

ZESZYTY NAUKOWE
WYDZIAŁU ETI POLITECHNIKI GDAŃSKIEJ

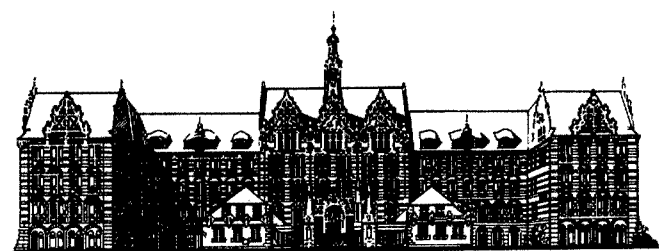
TECHNOLOGIE INFORMACYJNE



DISTRIBUTION STATEMENT A

Approved for Public Release
Distribution Unlimited

NR 4



1904 1945 2004/2005

JUBILEUSZ POLITECHNIKI W GDAŃSKU

BEST AVAILABLE COPY

20041206 053

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.			
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 23 August 2004	3. REPORT TYPE AND DATES COVERED Conference Proceedings, 18 May 2004	
4. TITLE AND SUBTITLE 2 nd Conference on Information Technology / Special Session on Homeland Security, Volume 4		5. FUNDING NUMBERS FA8655-04-1-5047	
6. AUTHOR(S) Conference Committee			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Gdansk University of Technology ul. Gabriela Narutowicza 11/12 80-952 Gdansk 80-952 Poland		8. Performing Organization Report Number	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 Box 14 FPO 09499-0014		10. SPONSORING/MONITORING AGENCY REPORT NUMBER CSP 04-5047	
11. SUPPLEMENTARY NOTES Volume 4, ISBN 83-917681-5-5 Copyright 2004 Wydział ETI Politechniki Gdanskiej Gdansk. Available from: Wydział ETI Politechniki Gdanskiej, Gdansk. The Department of Defense has permission to use for government purposes only. All other rights are reserved by the copyright holder.			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. (approval given by local Public Affairs Office)		12b. DISTRIBUTION CODE A	
ABSTRACT (Maximum 200 words) The Final Proceedings for 2 nd Conference on Information Technology / Special Session on Homeland Security, 16-18 May 2004 Formal methods in information engineering Electronic documents and digital libraries Autonomous robots in embeded systems Mobile and portable information systems Design and implementation of methodologies and technologies for information based products Dependability and security of information processing systems Development of IT infrastructure Special English Session on Homeland Security			
14. SUBJECT TERMS EOARD, Computational methods, C31, Computer network security			15. NUMBER OF PAGES 334
16. SECURITY CLASSIFICATION OF: a. Report UNCLASSIFIED b. Abstract UNCLASSIFIED c. This page UNCLASSIFIED	17. LIMITATIONS OF ABSTRACT UNCLASSIFIED	18a. NAME OF RESPONSIBLE PERSON Paul Losiewicz, Ph. D. 18b. TELEPHONE NUMBER (include area code) +44 20 7514 4474	

**ZESZYTY NAUKOWE
WYDZIAŁU ETI POLITECHNIKI GDAŃSKIEJ**

TECHNOLOGIE INFORMACYJNE

NR 4

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited



Gdańsk 2004

AQ F05-02-0396

KOMITET REDAKCYJNY

**Zeszytów Naukowych Wydziału ETI
Politechniki Gdańskiej**

Redaktorzy wydania

Józef Woźniak
Krzysztof Nowicki

Wydano za zgodą
DZIEKANA WYDZIAŁU ETI PG

W materiałach konferencyjnych zamieszczono
wyłącznie artykuły recenzowane

© Copyright by Wydział ETI Politechniki Gdańskiej
Gdańsk 2004

ISBN 83-917681-5-5

Druk: Zakład Poligrafii Politechniki Gdańskiej
ul. G. Narutowicza 11/12, 80-952 Gdańsk, tel. (0-58) 347 25 35

Technologie informacyjne (TI) umożliwiają implementację różnego typu systemów, szeroko stosowanych we współczesnym świecie. Nie oznacza to wcale, że są to już technologie dojrzałe, a ich wprowadzenie nie wiąże się z żadnym większym ryzykiem. Ogromne tempo rozwoju TI wynika z faktu, że korzyści z ich wykorzystania bądź nadzieja na osiągnięcie znacznych korzyści w przyszłości są ciągle duże. Na ogół konwergencja (integracja wielu istniejących rozwiązań) i synergia (generowanie nowych możliwości po integracji tych rozwiązań) są motorem ciągłego postępu.

II Konferencja Technologie Informacyjne jest zorganizowana przez Wydział Elektroniki, Telekomunikacji i Informatyki Politechniki Gdańskiej pod patronatem trzech Prezydentów Trójmiasta: Gdańska, Gdyni, Sopotu i J. M. Rektora PG w Roku Jubileuszowym Politechniki Gdańskiej, w rocznicę 100-lecia politechniki w Gdańsku. Odbywa się także już po fakcie ponownego włączenia się Polski do rodziny Krajów Europejskich. Stąd też większy zakres tematyczny tej konferencji, jak również częściowo międzynarodowy charakter. Znacznie powiększył się także skład Komitetu Programowego oraz liczba przyjętych do druku artykułów.

Tematy konferencji skupią się wokół 13 sesji tematycznych, w tym kilka sesji jest zorganizowanych przez uczestników konferencji. Są to: Homeland Security, Nietechniczne aspekty TI, Systemy mikroelektroniczne, Systemy radiowe. Tego typu sesje specjalistyczne chcemy rozwijać w przyszłości.

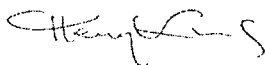
W pierwszym dniu Konferencji odbędzie się również specjalna sesja z okazji Światowego Dnia Telekomunikacji poświęcona problemom kształcenia, a wieczorem na bankiecie zostanie wręczony „Laur dla Pracodawcy” najlepszej firmie wybranej przez studentów, pracowników i absolwentów Wydziału ETI. W drugim dniu konferencji zorganizowane będą studenckie Targi Pracy. Odbędą się również dwie sesje plenarne, na których wybitni specjaliści przedstawią bardzo interesujące referaty:

1. Prof. Wojciech Szpankowski: *Ubiquitous Pattern Matching and its Applications*,
2. Prof. Józef Lubacz i prof. Andrzej Kraśniewski: *Ewolucja szkolnictwa wyższego w Europie i jej konsekwencje dla kształcenia w obszarze technik i technologii informacyjnych*.

Żałuję, że brakuje sesji poświęconej Twórcy Szkoły Mikrofal na Wydziale ETI PG Profesorowi Krzysztofowi Grabowskiemu. Niech więc wysoki poziom obrad oraz wyniki konferencji będą wyrazem podziękowania za Jego trud poświęcony rozwojowi naszego Wydziału.

Mam nadzieję, że obrady II Konferencji Technologie Informacyjne przyczynią się do lepszego wzajemnego poznania się wielu ekspertów z tej dziedziny. Wówczas dzięki skutkom konwergencji i synergii osiągniemy znacznie większy postęp we współpracy na rzecz zaspokajania potrzeb społeczeństwa informacyjnego.

Z podziękowaniem za współpracę



Dziekan Wydziału ETI
Politechnika Gdańska

Komitet Naukowy

Przewodniczący

Prof. Henryk Krawczyk (WETI PG)

Wiceprzewodniczący

Prof. Józef Woźniak (WETI PG)

Sekretarz

Dr Krzysztof Nowicki (WETI PG)

Członkowie:

Prof. Daniel Bem (WE PWr)
Prof. Michał Białko (PAN)
Prof. Zdzisław Bubnicki (PAN)
Prof. Andrzej Czyżewski (WETI PG)
Prof. Władysław Findeisen (PAN)
Prof. Tadeusz Galanc (WiZ PWr)
Prof. Janusz Górski (WETI PG)
Prof. Tomasz Imieliński (RU USA)
Prof. Zygmunt Kitowski (AMW)
Prof. Bogdan Kosmowski (WETI PG)
Prof. Marek Kubale (WETI PG)
Prof. Krzysztof Kuchciński (LIT, SE)
Prof. Józef Lubacz (WETI PW)
Prof. Stanisław F. Łęgowski (UW, USA)
Prof. Jerzy Mazur (WETI PG)
Prof. Maciej Niedźwiecki (WETI PG)
Prof. Antoni Nowakowski (WETI PG)
Prof. Tadeusz Orzechowski (WEALiE AGH)
Prof. Zdzisław Pawlak (PAN)
Prof. Krzysztof Pawlikowski (CU, NZ)
Prof. Michał Polowczyk (WETI PG)
Prof. Andrzej Ruciński (UNH, USA)
Prof. Dominik Rutkowski (WETI PG)
Prof. Jerzy Rutkowski (WAEil PŚI.)
Prof. Roman Salamon (WETI PG)
Prof. Wojciech Sobczak (WETI PG)
Prof. Andrzej Stepnowski (WETI PG)
Prof. Wojciech Szpankowski (PU, USA)
Prof. Jan Węglarz (PAN)
Prof. Stefan Węgrzyn (PAN)
Prof. Bogdan M. Wilamowski (AU, USA)
Prof. Bogdan Wiszniewski (WETI PG)
Prof. Wiesław Woliński (PAN)
Prof. Romuald Zielonko (WETI PG)
Prof. Marian Zientalski (WETI PG)
Prof. Jacek M. Żurada (UL, USA)

Recenzenci

Andrzej Czyżewski
Krzysztof Goczyła
Ewa Hermanowicz
Wojciech Jędruch
Jerzy Kaczmarek
Sylwester Kaczmarek
Renata Kalicka
Ryszard Katulski
Bogdan Kosmowski
Bożena Kostek
Henryk Krawczyk
Marek Kubale
Jerzy Mazur
Maciej Niedźwiecki
Antoni Nowakowski
Michał Polowczyk
Andrzej Ruciński
Dominik Rutkowski
Roman Rykaczewski
Wojciech Sobczak
Bogdan Wiszniewski
Józef Woźniak
Romuald Zielonko
Marian Zientalski

SPIS TREŚCI

ZAUTOMATYZOWANE SYSTEMY ŁĄCZNOŚCI RADIOWEJ KF Jan Cichy, Mariusz Błaś, Jerzy Pakieser	327
TŁUMIENIE SYGNAŁU RADIOWEGO W ŚRODOWISKU LEŚNYM Maciej J. Grzybowski	339
WSPÓŁCZESNA ŁĄCZNOŚĆ MORSKA Ryszard J. Katulski	347
ZMODYFIKOWANY MODEL COST 231 HATY DO PROJEKTOWANIA SYSTEMÓW DOSTĘPOWYCH Ryszard J. Katulski, Andrzej Kiedrowski	355
UKŁADY ANTENOWE MIMO W TELEKOMUNIKACJI BEZPRZEWODOWEJ Ryszard J. Katulski, Michał Mikołajski	363
INTERAKCJA ELEKTRODYNAMICZNA ANTENY Z UŻYTKOWNIKIEM TELEFONU KOMÓRKOWEGO Ryszard J. Katulski, Jarosław Sadowski	371
PROMIENIE LASERA NIEBIESKO-ZIELONEGO W ZASTOSOWANIU DO ŁĄCZNOŚCI Z OBIEKTAMI ZANURZONYMI Andrzej Kosiński	379
MONITORING RADIOWY OBIEKTÓW SZYBKO PRZEMIESZCZAJĄCYCH SIĘ NAD MORZEM Andrzej Kosiński, Bogusław Pilek	387
BADANIA SYMULACYJNE ALGORYTMU LS-DRTMA W ZASTOSOWANIACH ANTENOWYCH Michał Mikołajski	395
UWARUNKOWANIA PROPAGACYJNE W WSPÓŁCZESNEJ ŁĄCZNOŚCI MORSKIEJ Wiktor Pawłowski	403
KONCEPCJA SYSTEMU TRANSMISJI DANYCH DLA POTRZEB AKCJI RATOWNICZYCH Marek Przyborski	413
TRENDY ROZWOJOWE TRANSMISJI DANYCH WE WSPÓŁCZESNEJ MORSKIEJ RADIOKOMUNIKACJI WOJSKOWEJ Ryszard Studański	417
ZASTOSOWANIE KOMPUTEROWEJ REJESTRACJI MOWY W MONITORINGU RADIOWYM NA MORZU Marek Wesołowski, Mariusz Zieliński	425
THE IMPLEMENTATION OF A MICROSERVER Arnold Adamczyk	433
BADANIE JAKOŚCI TRANSMISJI MOWY W SIECIACH IP A. Czyżewski, A. Kaczmarek, P. Maziewski, P. Ody, P. Szczuko, Krzyżniewski, L. Smoleński, M. Zientalski	439
URZĄDZENIA FIRMY DGT W TECHNOLOGII VOIP I ICH INSTALACJA W PIŁOTAŻOWEJ SIECI Piotr Gajos, Sławomir Kwiatkowski	447
ODPORNOŚĆ NA AWARIE SIECI BEZSKALOWYCH Tomasz Gierszewski	455
DYNAMICZNE SIECI VLAN WYKORZYSTUJĄCE INFRASTRUKTURĘ KLUCZA PUBLICZNEGO PKI Grzegorz Górski	463
TRANSFER OF CHARGING INFORMATION BETWEEN EXCHANGES OF DIFFERENT NETWORK OPERATORS IN MULTI-PROVIDER ENVIRONMENT Samer Jamous	471
BADANIE WPLYWU RUCHU NIŻSZEGO PRIORYTETU NA JAKOŚĆ USŁUG O WYMAGANIACH CZASU RZECZYWISTEGO W SIECI IP QOS Sylwester Kaczmarek, Marcin Narloch	479

MODEL SYMULACYJNY I BADANIA SIECI MPLS Sylwester Kaczmarek, Krzysztof Nowak	487
WYBRANE WŁASNOŚCI PROBLEMU ROUTINGU ORAZ KOLOROWANIA ŚCIEŻEK W GRAFIE Adrian Kosowski	495
SYSTEM ZARZĄDZANIA RUCHEM W SIECI Z WĘZŁAMI DGT Leszek Mazur, Artur Stencel, Daniel R. Świeciak, Ryszard Weisbrodt	503
MODEL PRZEŻYWAŁNYCH SIECI IP NASTĘPNEJ GENERACJI Wojciech Molisz	511
METODA SZYBKIEGO WDRAŻANIA IPV6 PRZY WYKORZYSTANIU DHCPV6 Tomasz Mrugalski	519
PRZEŻYWAŁNOŚĆ POŁĄCZEŃ W FUNKCJI CZĘSTOŚCI AWARII ELEMENTÓW SIECI I ICH POŁOŻENIA GEOGRAFICZNEGO Jacek Rak	527
KIERUNKI ROZWOJU TECHNIK TRANSMISJI W WIELOUSŁUGOWYCH OPTYCZNO- MIEDZIANYCH SIECIACH DOSTĘPOWYCH Lech Smoleński	535
ZASADY WSPÓŁPRACY SIECI IPV4 I IPV6 ORAZ KIERUNKI MIGRACJI DO IPV6 Jacek Światowiak, Krzysztof Nowicki	543
ARCHITEKTURY SYSTEMU INFORMATYCZNEGO DLA PRZEDSIĘBIORSTW MAŁEJ I ŚREDNIEJ WIELKOŚCI Liliana Byczkowska-Lipińska, Marcin Cegielski	553
COMPUTEROWE ARCHIWIZOWANIE PRZEBIEGU PRACY GRUPOWEJ Beata Krawczyk-Brylka, Michał Piotrowski	561
PROINFO: SYSTEM ZARZĄDZANIA INFORMACJĄ O PRZEDSIĘWZIĘCIU PROGRAMISTYCZNYM Jerzy Nawrocki, Krzysztof Pospiech, Maciej Dorsz	569
ZMIANY OTOCZENIA BIZNESOWEGO A NAKŁAD PRACY W PROJEKCIE INFORMATYCZNYM – STUDIA PRZYPADKÓW Katarzyna Pardo, Wojciech E. Kozłowski	577
PROFIL KLIENTA W PROCESIE ZARZĄDZANIA RELACJAMI Z KLIENTAMI Justyna Stasienko	585
e-PR – PERSPEKTYWY PUBLIC RELATIONS W DOBIE SPOŁECZEŃSTWA INFORMACYJNEGO Sławomir Gawroński	591
JAKIEJ INFORMACJI CZŁOWIEK POTRZEBUJE NAPRAWDĘ Marek Hetmański	597
GLOBALNE SPOŁECZEŃSTWO INFORMACYJNE: TRZY TRENDY Kazimierz Krzysztofek	603
KONCEPCJA SYSTEMU ZDALNEGO NAUCZANIA W JAVIE Ilona Bluemke, Adrian Kuszewski, Konrad Łapienko	611
PROJEKT GRUPOWY W KSZTAŁCENIU STUDENTÓW INFORMATYKI - DOŚWIADCZENIA NA WYDZIALE ETI POLITECHNIKI GDAŃSKIEJ Anna Bobkowska, Maciej Piechówka, Stanisław Szejko	619
EDUKACJA W SPOŁECZEŃSTWIE INFORMACYJNYM Bogdan A. Galwas, Remigiusz J. Rak	627
ZASTOSOWANIE TECHNOLOGII INTERNETOWYCH W REALIZACJI SYSTEMÓW ZDALNEGO NAUCZANIA Krzysztof Gierłowski, Krzysztof Nowicki	635
MODEL SZKOLENIA "BLENDED LEARNING" Z WYKORZYSTANIEM PLATFORMY ORACLE I-LEARNING Anna Grabowska, Jerzy Durślewicz, Bartłomiej Antczak, Tomasz Skwarek	645
INTERNETOWY SYSTEM INFORMACYJNY O PROJEKTACH STUDENCKICH Jerzy Nawrocki, Renata Peliwo, Jacek Maza	653

Jan Cichy, Mariusz Błaś, Jerzy Pakieser

Ośrodek Badawczo-Rozwojowy Centrum Techniki Morskiej w Gdyni

ZAUTOMATYZOWANE SYSTEMY ŁĄCZNOŚCI RADIOWEJ KF

Streszczenie

Potrzeba dostosowywania systemów łączności radiowej KF SZ RP do wymagań stawianych przez Sojusz Północno-Atlantycki przyczyniła się do opracowania systemów łączności KF w OBR CTM spełniających założone cele. Powstałe systemy integrują modernizowane urządzenia z urządzeniami najnowszej generacji, zapewniając bezkonfliktową współpracę w danym systemie.

1. WSTĘP

Polska jako kraj członkowski Sojuszu Północno-Atlantyckiego została zobowiązana dostosować swoje siły zbrojne do struktur i systemów obowiązujących w państwach sojuszu. Zakres dostosowania, jak i związany z nim wymagany poziom nakładów finansowych decydują o rozłożeniu ich w czasie, co daje szansę krajowym instytucjom badawczym na znalezienie odbiorców swoich rozwiązań, konkurencyjnych ekonomicznie oraz technologicznie. Wśród powstałych opracowań będących odpowiedzią na potrzeby Sił Zbrojnych Rzeczypospolitej Polskiej, na szczególną uwagę zasługują systemy łączności radiowej KF.

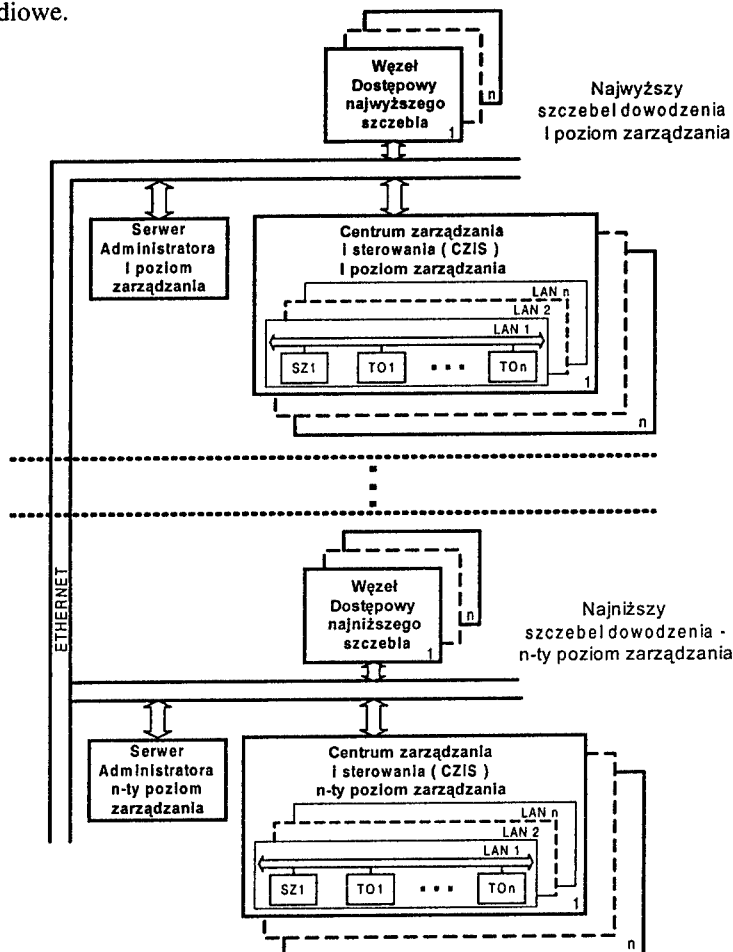
Jednym z fundamentów osiągnięcia powodzenia przez Siły Zbrojne RP podczas wykonywania zadań jest organizacja i eksploatacja efektywnego systemu łączności, który zabezpieczy proces kierowania wojskami na wszystkich wymaganych szczeblach dowodzenia. Wykorzystanie radiowego kanału KF na potrzeby takiego systemu łączności wydaje się najbardziej celowe, a dla krajów nie posiadających własnych systemów satelitar-nych, wręcz konieczne.

2. SYSTEM KIEROWANIA I ZDALNEGO STEROWANIA ŚRODKAMI RADIOWYMI KF DLA STANOWISK DOWODZENIA RÓŻNYCH SZCZEBLI WP

Efektywność wykorzystania posiadanego sprzętu radiowej KF przez SZ RP można zwiększyć poprzez wdrożenie systemu, który pozwoli na zdalne sterowanie i kierowanie nim oraz umożliwi optymalną eksploatację sprzętu radiowej starszej i nowszej generacji. Budowa takiego systemu pozwoli także na podniesienie bezpieczeństwa i obronności państwa.

Biorąc pod uwagę wymienione korzyści wydaje się celowe przeprowadzenie rozważań nad koncepcją nowoczesnego systemu kierowania i zdalnego sterowania środkami radiowymi KF dla stanowisk dowodzenia różnych szczebli.

Zasadniczym zadaniem systemu kierowania i zdalnego sterowania środkami radiowymi KF jest udostępnienie użytkownikom narzędzi oraz środków pozwalających im na wykorzystanie urządzeń radiowych [1]. Podstawowymi elementami systemu realizującymi ww. zadanie są: serwer administratora (SA), centrum zarządzania i sterowania (CZiS) oraz węzeł dostępowy KF (WD). CZiS stanowi strukturę organizacyjną systemu utworzoną przez terminale operatorskie (TO) wraz z serwerem zarządzania (SZ). Przedstawione elementy wraz z kierowanymi przez nie środkami radiowymi tworzą całkowicie nowy system zarządzania łącznością radiową KF. Przykładową konfigurację systemu zarządzania łącznością radiową KF przedstawia rys. 1. Jego organizacja oparta jest na obowiązującej strukturze dowodzenia i obejmuje szereg poziomów zarządzania o określonej hierarchii kompetencji oraz odpowiedzialności. Każdy poziom zarządzania zabezpiecza funkcjonowanie jednego lub kilku podsystemów łączności radiowej KF w zależności od potrzeb systemu dowodzenia. Komunikacja i wymiana danych systemowych pomiędzy elementami systemu odbywa się przy wykorzystaniu sieci pracującej w standardzie Ethernet. Jako medium transmisyjne można wykorzystywać łącza przewodowe, światłowodowe, radioliniowe lub radiowe.



Rys. 1. Przykładowa konfiguracja systemu zarządzania łącznością radiową KF

2.1. Przeznaczenie i zadania serwera administratora

Serwer administratora (SA) jest przeznaczony do pracy dla szefa łączności odpowiedniego szczebla dowodzenia (np. najwyższego, pośredniego oraz podstawowego). Budowa SA została oparta na komputerze klasy PC wyposażonym w odpowiednie oprogramowanie oraz interfejs sieciowy zgodny ze standardem Ethernet. Konfiguracja sprzętowa komputera gwarantuje efektywne wykorzystanie platformy systemowej Windows 2000. Oprogramowanie SA umożliwia szefowi łączności: planowanie, organizowanie i nadzorowanie funkcjonowania systemu na własnym oraz podległym poziomie zarządzania. Na podstawie otrzymanych baz danych, szef łączności wykorzystując SA, rozdziela parametry pracy elementów systemu na potrzeby organizacji dowodzenia własnego oraz podległego szczebla. Biorąc pod uwagę zadania wykonywane przez podległe jednostki wojskowe, opracowuje bazy danych niezbędne do organizacji systemu łączności na własnym oraz podległym poziomie zarządzania.

Oprogramowanie SA zapewnia:

- przeprowadzenie procedury autoryzacji dostępu operatora do SA;
- pobieranie baz danych systemowych z SA wyższego poziomu zarządzania;
- wspomaganie procesów: planowania i nadzorowania działania systemu zarządzania łącznością radiową KF na poziomie podległym oraz własnym;
- archiwizowanie baz danych systemowych koniecznych do funkcjonowania systemu na własnym oraz podległym poziomie;
- udostępnianie podległym SA i SZ niezbędnych do pracy baz danych;
- wykonanie i przekazywanie do SA wyższego poziomu zarządzania raportów o stanie podległej części systemu zarządzania łącznością radiową KF.

2.2. Przeznaczenie i zadania serwera zarządzania

Serwer zarządzania (SZ) jest przeznaczony dla szefa węzła łączności zabezpieczającego funkcjonowanie stanowiska dowodzenia. Budowa SZ została oparta na komputerze klasy PC wyposażonym w odpowiednie oprogramowanie oraz interfejsy sieciowe zgodne ze standardem Ethernet. Konfiguracja sprzętowa komputera gwarantuje efektywne wykorzystanie platformy systemowej Windows 2000. Na węźle łączności w ramach CZiS może być eksploatowany tylko jeden SZ. Podstawowym zadaniem SZ jest zapewnienie kierowania i zdalnego sterowania środkami radiowymi KF pracującymi na potrzeby danego węzła łączności oraz zarządzanie dostępem do łączności radiowej KF osób funkcyjnych uprawnionych do prowadzenia korespondencji radiowej (dowódcy właściwego szczebla).

Oprogramowanie SZ zapewnia w systemie:

- przeprowadzenie procedury autoryzacji dostępu operatora do SZ;
- pobieranie z SA baz danych systemowych;
- wspomaganie procesów: planowania i eksploatowania łączności radiowej KF na potrzeby stanowiska dowodzenia;
- podporządkowanie podległych TO i WD zgodnie z danymi otrzymanymi z SA;
- wykonanie i przekazywanie do przełożonego SA raportów o stanie podległej części systemu zarządzania łącznością radiową KF;
- przydzielenie odpowiednich środków radiowych do realizacji zadań transmisyjnych;
- dystrybucję transmitowanych przez system wiadomości i sygnałów sterujących;
- zwiększenie możliwości dostępu abonentów systemu do wykorzystywanych środków radiowych;

- realizację funkcji nadzoru wraz z zobrazowaniem stanu pracy podległych TO i WD;
- alarmowanie SA o sytuacjach awaryjnych w podległym CZiS.

2.3. Przeznaczenie i zadania terminala operatorskiego

Terminal operatorski (TO) zapewnia osobom funkcyjnym poszczególnych szczebli dowodzenia komunikowanie się przy wykorzystaniu łączy radiowych zestawianych i zarządzanych przez system. Zadania TO w systemie może wykonywać dowolny komputer klasy PC wykorzystujący platformę systemową Windows 2000. Praca TO opiera się na wykorzystaniu dowolnego klienta poczty elektronicznej pracującego w środowisku Windows 2000.

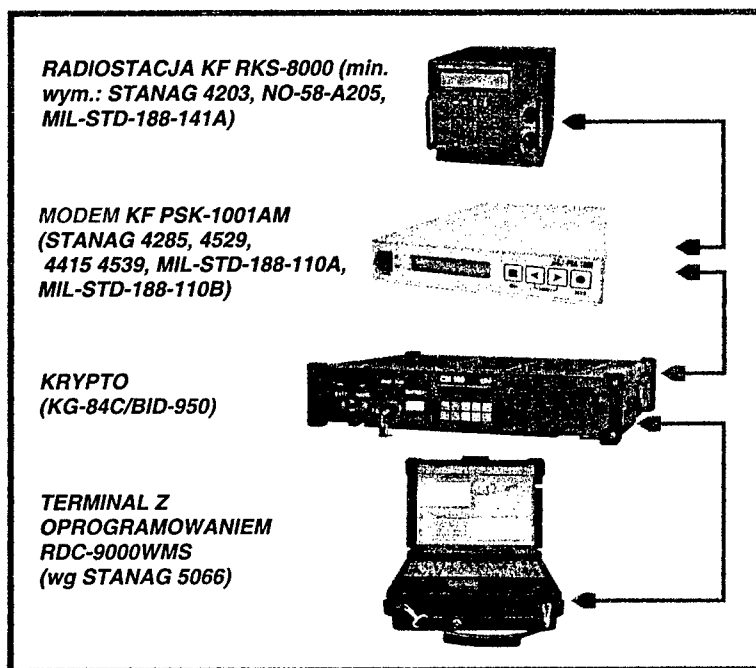
2.4. Przeznaczenie i zadania węzła dostępowego KF

Zapewnienie zautomatyzowanej transmisji w sieciach KF realizuje węzeł dostępowy stanowiący integralny element węzłów łączności dowolnego szczebla dowodzenia. Szczegółowe wymagania dotyczące budowy węzła dostępowego KF (WD) precyzuje uzgodnienie STANAG 5066. OBR CTM jest autorem oprogramowania RDC-9000WMS stanowiącego implementację wymagań STANAG-u 5066, a także modemu KF i radiostacji KF, które są integralnymi elementami WD (rys. 2).

Węzeł dostępowy KF jest elementem, który umożliwia wykorzystanie w systemie środków radiowych starszej i nowszej generacji. Jeden WD zabezpiecza eksploatację w systemie jednego środka radiowego, zapewnia sterowanie nim w zakresie przewidzianym przez producenta oraz wykorzystuje go do transmisji danych drogą radiową. Budowa WD jest oparta na komputerze klasy PC wyposażonym w odpowiednie oprogramowanie oraz zestaw interfejsów do pracy z urządzeniami radiowymi i dodatkowym wyposażeniem WD. Konfiguracja sprzętowa komputera gwarantuje efektywne wykorzystanie platformy systemowej Windows 2000. Sygnały sterujące, transmitowane do WD z SZ przy wykorzystaniu sieci, zgodnie ze standardem Ethernet, są zamieniane na odpowiednie komendy sterujące interpretowane przez sprzęt radiowy. Statusy pracy urządzeń radiowych, po przetworzeniu w WD, zostaną przesłane do SZ. Przyjęta koncepcja WD pozwala na wykorzystanie go jako węzła KF w rozumieniu uzgodnienia STANAG 5066.

Oprogramowanie WD zapewnia w systemie:

- przeprowadzenie procedury autoryzacji dostępu operatora do WD;
- pobieranie z SZ danych do pracy urządzeń radiowych oraz dodatkowego wyposażenia WD;
- transmitowanie drogą radiową wiadomości przekazywanych z SZ;
- konwersje sygnałów sterujących otrzymywanych z SZ na komendy sterujące interpretowalne przez urządzenia radiowe (np. modemy KF);
- konwersje statusów pracy urządzeń radiowych (np. modemów KF) na sygnały sterujące interpretowalne przez SZ;
- realizację funkcji diagnostycznych wyposażenia WD oraz środków radiowych;
- alarmowanie SZ o sytuacjach awaryjnych na WD.



Rys. 2. Węzeł dostępowy KF produkcji OBR CTM.

2.5. Charakterystyka oprogramowania RDC-9000WMS

Oprogramowanie RDC-9000WMS zostało opracowane i wykonane w myśl wymagań STANAG-u 5066 v 1.2. Spełnia ono normy określone przez wojskowe biuro standaryzacji (Podkomitet ds. Sieci Łączności) w NATO C3 posiada następujące możliwości:

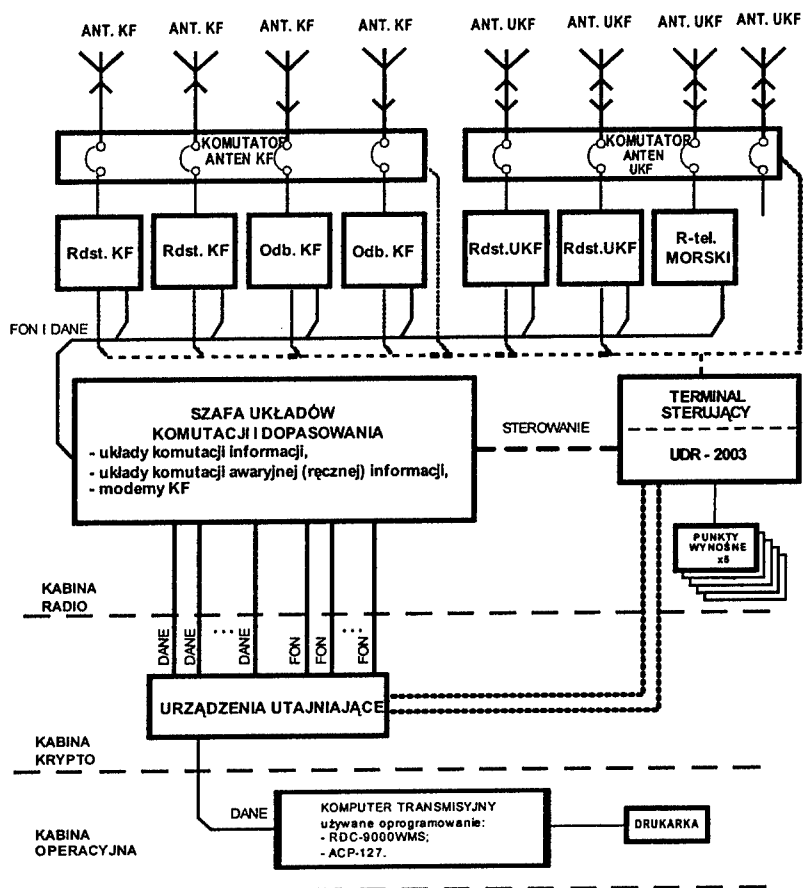
- Obsługuje systemy operacyjne:
 - Windows NT.x.;
 - Windows 2000;
- Posiada interfejs TCP/IP klient/serwer,
- Współpracuje z serwerem pocztowym (SMTP i POP3);
- Współpracuje z dowolnym klientem poczty elektronicznej obsługującym ww. protokoły np.:
 - MS Outlook;
 - Lotus Notes;
 - Eudora;
- Posiada zaimplementowanych klientów: HMTP, SMTP i IP.

Węzeł KF, wykonany w OBR CTM na bazie urządzeń i oprogramowania własnego projektu i produkcji, jest w pełni interoperacyjny ze sprzętem eksploatowanym przez wojska NATO, co zostało potwierdzone testami w Niemczech na ćwiczeniach COMBINED ENDEAVOR 2003.

3. SYSTEM ZARZĄDZANIA RADIOWYM CENTRUM NADAWCZO-ODBIORCZYM OKRĘTU

Zautomatyzowane centrum nadawczo-odbiorcze okrętu (ZCN-O) [2] zapewnia w szerokim zakresie interoperacyjność z systemami łączności sił morskich państw NATO oraz integruje elementy podsystemu łączności, znacznie zwiększając przy tym, stopień jego automatyzacji. System spełnia wymagania NATO (STANAG 4529, 4285, 4539, 4415, MIL-STD-188-110A/B). Współpraca z narodowym systemem łączności jest zapewniona przez obsługę krajowych urządzeń transmisyjnych i utajniających.

W wyniku przyjęcia założenia pełnej podatności na sterowanie wszystkimi parametrami i funkcjami urządzeń łączności na okręcie, powstał zaawansowany system informatyczny zapewniający jednocześnie w pełni elastyczną konfigurację całego systemu ZCN-O.



Rys. 3. Układ funkcjonalny systemu ZCN-O

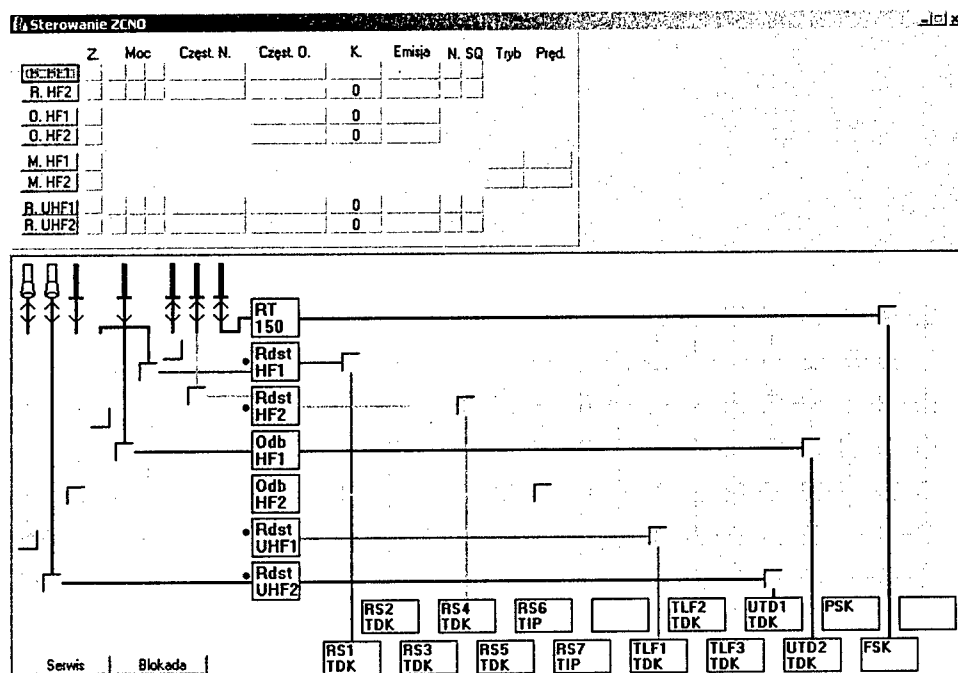
Opracowany i wykonany system zapewnia:

- sterowanie komutatorami anten KF i UKF (typ KHF-1621 i KVF-1621 produkcji OBR CTM),
- sterowanie urządzeniami radiowymi KF i UKF (np. R&S, RKS-8000 produkcji OBR CTM),

- sterowanie modemami radiowymi KF (np. typ PSK-1001AM produkcji OBR CTM),
- sterowanie komutatorami danych (typ KTD-1621 produkcji OBR CTM),
- sterowanie urządzeniem dostępu radiowego (np. typ UDR-2003/UR5 produkcji OBR CTM).

Układ funkcjonalny systemu zarządzania przedstawia rys. 3.

System daje operatorowi możliwość sterowania środkami radiowymi za pomocą oprogramowania systemowego, zainstalowanego na terminalu sterującym, ze swojego miejsca pracy w pełnym zakresie przewidzianym przez producenta sprzętu radiowego. Terminal sterujący (TS) ZCN-O odpowiada za komunikację systemu z operatorem poprzez interfejs graficzny (rys. 4.) wyświetlany na monitorze dotykowym. Operator może, dzięki zainstalowanemu na TS oprogramowaniu, przeprowadzać komutację sygnałów informacyjnych (dane, fon) z dostępnymi kanałami radiowymi, komutację anten z odpowiednimi środkami radiowymi oraz sterowanie parametrami obsługiwanych urządzeń radiowych (radiostacji KF i UKF, odbiorników KF, modemów KF). Komputer TS, będący jednostką centralną, pracuje pod kontrolą systemu Windows NT 4.x. Aktualnie dostępne są interfejsy przeznaczone do sterowania środkami radiowymi firmy Rohde & Schwarz i OBR CTM, jednakże system ZCN-O pozwala na implementację interfejsu dowolnego producenta środków radiowych.



Rys. 4. Interfejs graficzny systemu ZCN-O na Terminalu Sterującym

Diagnozowanie oraz predykcję stanów awaryjnych systemu oparto na analizie nadchodzących statusów z poszczególnych elementów systemu. Mimo zaawansowanych form zabezpieczeń przed awariami oraz dla zapewnienia wyższej gotowości operacyjnej systemu, przewidziano jednocześnie możliwość ręcznej awaryjnej komutacji anten i urządzeń informacyjnych z urządzeniami łączności.

Transmisja danych w kanale krótkofalowym realizowana jest za pośrednictwem komputera transmisyjnego (rys. 3), przy wykorzystaniu modemu radiowego KF typ PSK-1001AM. Modem spełnia wymagania norm NATO ujęte w uzgodnieniach STANAG 4285, 4529, 4539 i 4415 oraz MIL-STD-188-110A/B. Polecenia zdalnego sterowania modemem są zgodne ze STANAG 5066, aneks E. Prędkość transmisji danych, dzięki zastosowaniu różnych metod modulacji (FSK, PSK, QAM), zawiera się od 50 do 19200 bitów/s. Modem gwarantuje współpracę z szeroką gamą sprzętu radiowego KF, dzięki regulacji poziomu sygnału wyjściowego w zakresie $-20\text{dB} \div +6\text{dB}$, zapewniając tym samym dużą elastyczność systemu ZCN-O. Komputer transmisyjny może wykonywać zadania, oprócz transmisji danych w formacie ACP-127 i poczty elektronicznej, serwera lokalnej sieci komputerowej TCP/IP przy wykorzystaniu oprogramowania RDC-9000WMS.

4. SYSTEM STEROWANIA ŚRODKAMI ŁĄCZNOŚCI RADIOWEJ I JEDNOCZESNEJ WYMIANY INFORMACJI FONICZNEJ ZA POŚREDNICTWEM POJEDYNCZEGO ŁĄCZA TELEFONICZNEGO.

Śmiało należy postawić tezę, że we współczesnym świecie występuje tendencja do integrowania środków łączności w spójny, łatwy do centralnego sterowania, system. Rozproszone w terenie środki łączności radiowej (radiostacje KF i UKF) łatwo jest połączyć z centrum dowodzenia za pomocą istniejących standardowych łącz telefonicznych. W ramach jednego systemu może występować kilka centrów dowodzenia o różnych priorytetach dostępu do środków łączności radiowej. Każde centrum może się składać z kilku stanowisk dowodzenia, również o zróżnicowanych priorytetach.

Na podstawie powyższych założeń w OBR CTM opracowany został System Zarządzania Środkami Radiowymi dla Potrzeb Ratownictwa Wojskowego na Morzu, zwany SZR-2010, przeznaczony do zarządzania i sterowania radiostacjami w dyspozycji ratownictwa MW.

4.1. Budowa urządzeń systemu

Ze względu na wymaganą uniwersalność zastosowano urządzenia o budowie pakietowej. Pakiety zbudowane są w oparciu o programowalne układy elektroniczne i pozwalają na dostosowanie systemu do konkretnych zastosowań i konkretnych typów środków łączności radiowej.

Końcówkami systemu są:

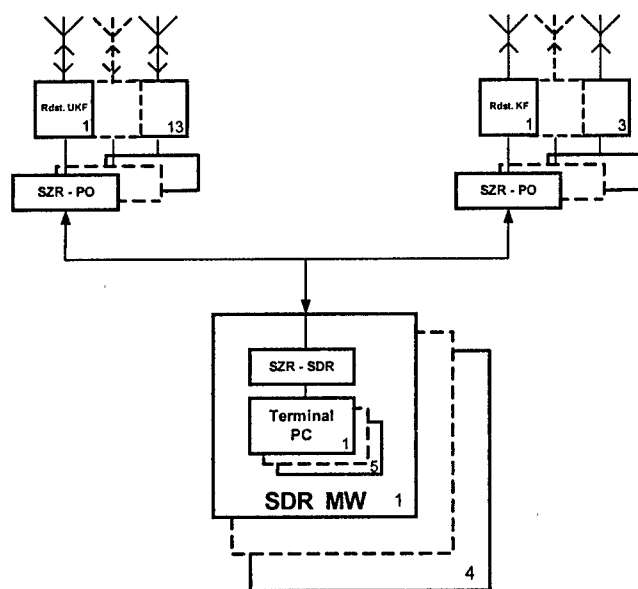
- centra dowodzenia w postaci połączonych w sieć LAN terminali PC z zainstalowanym oprogramowaniem zarządzającym, pracującym pod kontrolą systemu WINDOWS NT.x. (terminale wyposażone są w układy rozmówne do prowadzenia łączności fonicznej),
- dowolne radiostacje wyposażone w możliwość zdalnego sterowania za pomocą interfejsu RS232 (w przypadku radiostacji starszych typów istnieje możliwość wykonania stopni sterujących umożliwiających zastosowanie ich w systemie np. radiostację R-140 wyposażono w opracowane przez OBR CTM bloki sterujące: USN-140 oraz USO-155).

Kluczowym elementem systemu jest pakiet komunikacyjny BT, który pełni funkcję sterowanego mikroprocesorem modemu i sumatora/separatora sygnałów. Do budowy pakietu BT wykorzystano nowoczesne układy procesorów DSP na bazie, których wykonany moduł spełnia jednocześnie funkcje układu transmisji danych i przesyłania sygnałów akustycznych (mowy).

Wszystkie operacje wykonywane są przez procesor sygnałowy zestawiający dwa dwukierunkowe tory połączeniowe: akustyczny i sterujący. Tor akustyczny służy do przesyłania sygnałów mowy pomiędzy radiostacją i układem rozmównym stanowiska dowodzenia. Tor sterujący służy do przeniesienia komend sterujących i komunikatów informacyjnych pomiędzy komputerem a radiostacją. Całością operacji wykonywanych przez pakiet steruje mikroprocesor z własnym programem pracy. Pakiet BT umożliwia także dopasowanie do impedancji linii telefonicznej oraz automatyczną korekcję poziomu sygnałów sterujących.

4.2. Oprogramowanie zarządzające systemem

System umożliwia konfigurowanie łączności w taki sposób, aby jak najefektywniej zarządzać i koordynować pracę sił ratowniczych biorących udział w akcji poszukiwawczo-ratowniczej [3]. Rysunek 5 obrazuje schematyczne połączenia pomiędzy odległymi punktami w systemie.



Rys. 5. Schemat organizacji systemu łączności dla potrzeb ratownictwa morskiego typ SZR - 2010.

Koordynacja akcją poszukiwawczo-ratowniczą jest prowadzona z SDR MW (Stowisko Dowodzenia Ratownictwem MW), czyli centrum koordynacji lub dowodzenia akcją ratowniczą.

Każde z SDR MW może być zlokalizowane w dowolnym miejscu, pod warunkiem możliwości zestawienia odpowiedniej ilości łączy telefonicznych do obiektów łączności radiowej. SDR MW wyposażone jest w zestaw układów sterujących poszczególnymi radiostacjami przeznaczonymi do dowodzenia akcją poszukiwawczo-ratowniczą. Zestaw sterujący, nazwany sterownikiem komunikacyjnym SZR-SDR, jest obsługiwany za pomocą komputerów klasy PC z zainstalowanym programem koordynacji łączności radiowej. Stanowisko dowodzenia steruje i zarządza obiektami łączności radiowej znajdującymi się w różnych punktach całego wybrzeża morskiego RP. Obiekty łączności radiowej wyposażone są w układ sterujący bezpośrednio radiostacją, który nazwano sterownikiem obiekto- wym SZR-PO. Zestawy systemu SZR-2010: sterujący SZR-SDR i wykonawczy SZR-PO,

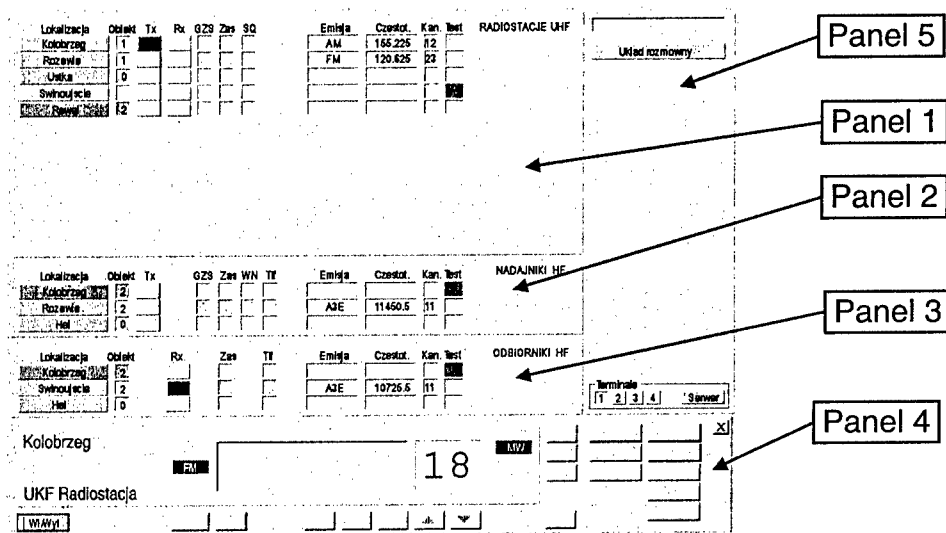
jako zasadnicze elementy systemu, są przygotowane do komunikowania się przez standardowe łącza telefoniczne. Jak przedstawiono na rysunku nr 5, każdy obiekt łączności radiowej jest dostępny dla wszystkich stanowisk dowodzenia SDR MW [4].

Oprogramowanie zarządzające, zwane Konsolą Sterującą pozwala na:

- zobrazowanie stanu środków łączności dostępnych w systemie,
- zestawianie połączeń z wybranymi środkami łączności,
- sterowanie parametrami pracy środków łączności,
- prowadzenie łączności fonicznej.

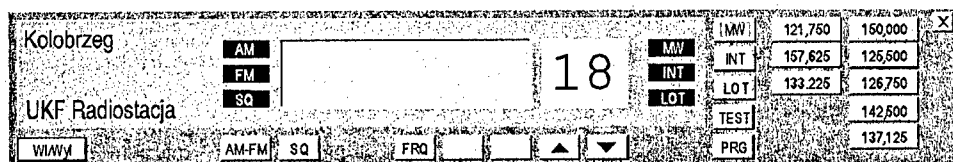
Zakres sterowania parametrami radiostacji ograniczony jest jedynie możliwościami do zdalnej obsługi w zależności od jej typu.

Zarządzanie systemem SZR-2010 odbywa się z terminali PC będących stanowiskami roboczymi w centrach dowodzenia za pomocą programu komputerowego o nazwie: „Konsola Sterująca SZR MW”. W oknie głównym programu dokonywane są wszystkie operacje związane z nawiązywaniem łączności radiowej i sterowaniem poszczególnymi radiostacjami. Ekran podzielony jest na pięć podstawowych paneli (patrz rys. 6). Panele od 1 do 3 opisują lokalizację, stan i wysterowanie poszczególnych środków łączności radiowej. Każdy środek łączności radiowej opisany jest w jednym wierszu odpowiedniego panelu.



Rys. 6. Wygląd okna głównego Konsoli Sterującej.

Panel 1 opisuje stany radiostacji UKF, Panel 2 – stany nadajników KF, a Panel 3 – stany odbiorników KF. Panel 4 służy do sterowania wybranym środkiem łączności radiowej i w zależności od jego typu, przyjmuje odpowiadającą mu postać – rys. 7 pokazuje przykładowy jego wygląd dla radiostacji UKF. Panel 5 zobrazowuje pozostałe parametry systemu, takie jak komunikaty o błędach, stan układu rozmównego, a także numer terminala w sieci LAN.



Rys. 7. Wygląd przykładowego pulpitu sterowania włączonej radiostacji UKF.

System zarządzania ratownictwem MW umożliwia sterowanie akcją ratowniczą na morzu z założonej ilości stanowisk dowodzenia, z uwzględnieniem nadanych priorytetów, przy czym w przypadku każdego ze stanowisk istnieje w skrajnym przypadku możliwość sterowania i zarządzania wszystkimi dostępnymi obiektami łączności radiowej (radiostacjami).

6. ZAKOŃCZENIE

Przedstawione systemy zarządzania łącznością radiową KF na potrzeby dowodzenia i zarządzania w SZ RP charakteryzują się następującymi właściwościami:

- możliwością wykorzystania w jednym systemie sprzętu radiowego starszej i nowszej generacji;
- adaptacyjną konfiguracją systemu w zależności od sytuacji militarnej lub politycznej;
- pełną spójnością z każdym wariantem systemu dowodzenia;
- brakiem ograniczeń, co do możliwości rozbudowy systemu w terenie;
- elastycznością, co do ilości wykorzystywanego sprzętu i poziomów zarządzania;
- wykorzystaniem, jako platformy systemowej, Windows 2000;
- możliwością wykorzystania w większości przypadków, komercyjnych komputerów klasy PC spełniających wymagania platformy systemowej Windows 2000;
- wysoką skrytością, pewnością i terminowością transmisji wiadomości oraz sygnałów sterujących;
- możliwością łatwego wzbogacania o nowy sprzęt i funkcje.

Wymienione cechy systemów pozwalają zaliczyć je do rozwiązań nowoczesnych. Należy również oczekiwać, że oferowany szeroki wachlarz możliwości pozwoli spełnić wymagania i aktualne oraz perspektywiczne potrzeby SZ RP.

BIBLIOGRAFIA

- [1] Opracowanie OBR CTM: „*Projekt wstępny zautomatyzowanego systemu kierowania łącznością radiową i zdalnego sterowania środkami radiowymi dla potrzeb brzegowych stanowisk dowodzenia*”, Oddział RC, Gdynia 2002.
- [2] Opracowanie OBR CTM: „*Koncepcja zautomatyzowanego centrum nadawczo-odbiorczego okrętu*”, Oddział RC, Gdynia 2000.
- [3] Kmdr dypl. Siudut L., kmdr ppor. mgr inż. Affelt J.: „*Projekt koncepcyjny systemu zarządzania środkami radiowymi dla potrzeb ratownictwa wojskowego na morzu*”, Oddział RC, Gdynia 2000.
- [4] Opracowanie OBR CTM: „*Projekt wstępny na system zarządzania środkami radiowymi dla potrzeb ratownictwa wojskowego na morzu*”, Oddział RC, Gdynia 2000.
- [5] Konieczny K. „*Potrzeby systemu dowodzenia, a możliwości taktycznej sieci telekomunikacyjnej*”, AON, Warszawa 1997.
- [6] STANAG 5066 (wersja 1.2) – *Wymagania na transmisję danych w kanale krótkofalowym (KF)*.

AUTOMATIED HF RADIO-COMMUNICATION SYSTEMS

Summary

The need for adaptation of HF radio communication systems in the Polish armed forces to requirements of NATO was reason for development of HF communication systems in R&D Marine Technology Centre, which perform specified aims. These systems integrate present arrangements with new generation equipment providing optimal cooperation in systems.

Maciej J. Grzybowski

Samodzielna Pracownia KEM we Wrocławiu, Instytut Łączności

TŁUMIENIE SYGNAŁU RADIOWEGO W ŚRODOWISKU LEŚNYM

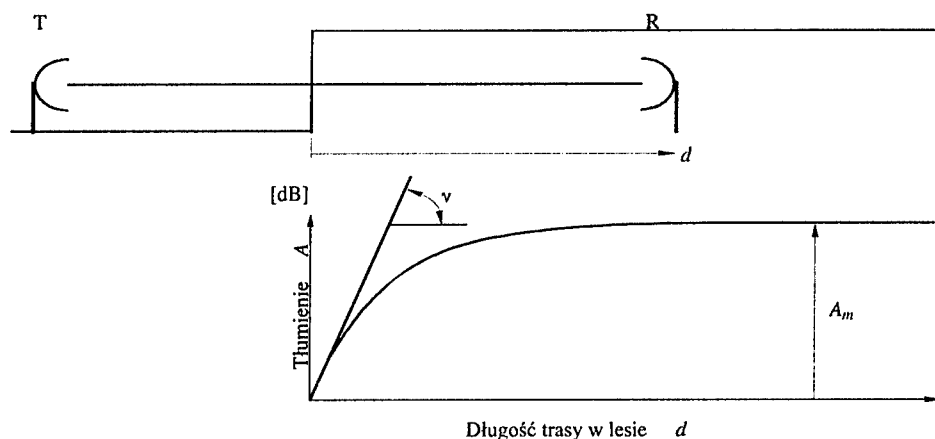
Streszczenie

W referacie przedstawiono wyniki badań tłumienia sygnałów radiowych o częstotliwościach zakresu VHF i UKF w aspekcie opracowanego w Instytucie Łączności modelu propagacji fal w środowiskach leśnych używanego do zasięgowych obliczeń komputerowych. Użycie tego modelu do obszarowych obliczeń propagacyjnych z wykorzystaniem cyfrowego modelu terenu pozwala na wprowadzenie poprawek umożliwiających uwzględnianie tłumienności wprowadzanej przez las. Założono, że dominującym efektem wpływającym na tłumienie fali wnikającej w głąb lasu jest dyfrakcja fali rozchodzącej się ponad nim i po jego obrzeżu. Rozważaniom podlega przyziemna trasa radiowa z jednym terminalem umieszczonym w głębi lasu. Badania potwierdziły przewidywany charakter tłumienia w obszarze lasu.

1. WSTĘP

Środowisko leśne wprowadza dość istotne tłumienie fal radiowych, szczególnie w zakresach VHF i UHF. Tłumienie to związane jest oddziaływaniem drzew i poszycia na proces rozchodzenia się fal elektromagnetycznych. Analiza propagacji fal radiowych w środowisku leśnym prowadzi do wniosku, że tłumienie lasu (ściśle związane ze strukturą lasu i porą roku) silnie rośnie w miarę zwiększania częstotliwości.

Z uwagi na skomplikowaną i w zasadzie niepowtarzalną strukturę różnych terenów leśnych Międzynarodowy Związek Telekomunikacyjny w Zaleceniu ITU-R PN.833 [1] podał do wiadomości wyniki pomiarów średnich wartości tłumienia fal radiowych w środowisku roślinnym – w przeliczeniu na dB/m. Zalecenie to było w ostatnich latach kilkakrotnie modyfikowane. Jednak w każdej postaci tego zalecenia model propagacyjny lasu objawiał się wzrostem tłumienia począwszy od skraju lasu do pewnej odległości, gdzie zaobserwowano pewną, ustaloną wartość tłumienia. Istotne było też, że tłumienie fal radiowych w niższych zakresach częstotliwości (do 1000 MHz) było uzależnione od polaryzacji fali – silniejsze było dla polaryzacji pionowej.

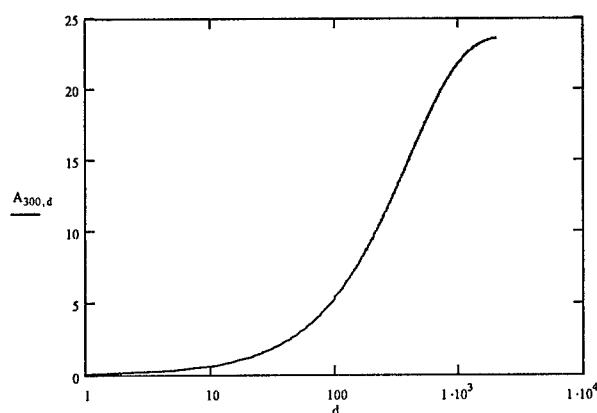
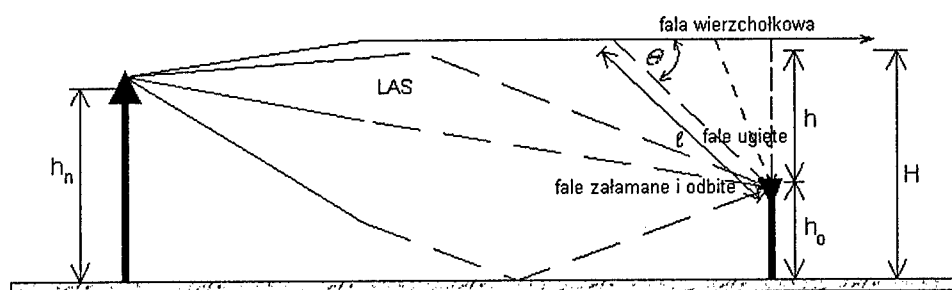


Rys. 1. Ilustracja modelu ITU rozchodzenia się fal radiowych w lesie [1]

Szczegółowe obserwacje rozkładu natężenia pola sygnału odbieranego w lesie pozwoliły na wniosek, że fala radiowa rozchodzi się głównie nad wierzchołkami drzew (tzw. fala wierzchołkowa) a następnie wnika w las na zasadzie dyfrakcji – w kierunku anteny odbiorczej. Pewna część fali przechodzi przez strukturę leśną bezpośrednio – od skraju lasu do anteny odbiorczej – ulegając po drodze odbiciu i załamaniu na granicy warstw las-powietrze i las-ziemia. Główne znaczenie dla odbioru będzie miała ta część fali radiowej, która jest najmniej wytłumiona na trasie. Identyczna sytuacja (w odwrotnej kolejności) tworzy się, gdy w przypadku, gdy wewnątrz kompleksu leśnego znajduje się antena nadawcza. Szczegółowe rozważania na ten temat podjęto w USA w latach 60. i 70. ubiegłego stulecia – prawdopodobnie w związku z wojną wietnamską. Sporządzono wtedy warstwowo-półprzestrzenny model lasu i teoretycznie analizowano przebieg fali radiowej [2]. Niestety, model ten okazał się niezbyt przydatny do wykonywania komputerowych obliczeń obszarowych a ponadto wymaga znajomości elektrycznych parametrów lasu.

W latach pięćdziesiątych w dobie gwałtownie rozwijającej się radiokomunikacji ruchomej lądowej istotne okazało się zapewnienie łączności „zawsze i wszędzie”. Uwzględnianie wpływu terenów zalesionych w predykcji rozkładu natężenia pola zaczęło nabierać szczególnego znaczenia przy określaniu zasięgów łączności dla rozmieszczonych poza miastami stacji bazowych systemów komórkowych. Powstało wówczas wiele modeli rozchodzenia się fal radiowych, jednak ze względu na dążenie autorów do możliwie wiernego zamodelowania środowiska propagacyjnego są one praktycznie nieprzydatne do komputerowych obliczeń obszarowych, gdzie wypadkowy rozkład natężenia pola tworzy po dokonaniu dużej liczby obliczeń w wybranych punktach obszaru otaczającego anteną nadawczą.

Z tego względu, dla potrzeb szacunkowych obliczeń obszarowych, w Instytucie Łączności utworzono odrębny model propagacyjny kompleksu leśnego [3]. Założono tutaj, że las stanowi jednolitą warstwę tłumiącą a tłumienie wnoszone przez środowisko w głąbi lasu jest stałe. Fala radiowa dochodząca do anteny odbiorczej podlega dyfrakcji na wierzchołkach drzew oraz tłumieniu w warstwie leśnej. Wypadkowe tłumienie jest zależne od sposobu (kąta) wnikania fal radiowych w warstwę leśną [4] oraz głębokości warstwy tłumiącej.



Rys. 2. Ilustracja modelu IŁ rozchodzenia się fal radiowych w lesie, tłumienie A [dB] w funkcji odległości d [m]; przykład dla $f=300$ MHz, polaryzacja pozioma [5].

Rozpatrywany model IŁ [3, 5] łatwo może być wykorzystywany do komputerowych obliczeń propagacyjnych, szczególnie przy użyciu cyfrowego modelu terenu DTM. Musiał być jednak potwierdzony na drodze pomiarowej.

2. WYNIKI BADAŃ PROPAGACYJNYCH

Pomiary propagacyjne są w środowisku leśnym dokonywane stosunkowo rzadko i tylko nieliczne pozycje literaturowe opisują ich wyniki. Jak dotąd związane jest to było prawdopodobnie z dużymi trudnościami napotkanymi podczas ich wykonywania ze względu na trudną dostępność do założonych punktów pomiarowych umieszczonych wewnątrz kompleksów leśnych. Dopiero w ostatnich latach niezbyt kosztowne przyrządy pomiarowe do wyznaczania natężenia pola elektrycznego uległy takiej miniaturyzacji, że można je swobodnie transportować w gęstwinie leśnej.

Instytut Łączności mógł dokonać badań tłumienia fali radiowej w lesie dopiero po uzyskaniu niezbędnego sprzętu pomiarowego. Badania te przeprowadzono w takich zakresach częstotliwości, gdzie możliwe jest wykorzystanie sygnałów emitowanych w systemach rozświecznych (radiofonicznych i telewizyjnych) oraz systemach operatorów komórkowych. Bez ponoszenia specjalnych nakładów na sprzęt nadawczy możliwe stało się wykonanie pomiarów w pasmach 100 MHz, 500 MHz, 900 MHz i 1800 MHz gdzie pracowały łatwe do zlokalizowania nadajniki stosownych sygnałów. Założono wtedy, że pomiary powinny być wykonane w różnych porach roku ze względu na różnice w ulistwieniu drzew. Oczekiwano, że wyniki pomiarów umożliwią albo potwierdzenie słuszności wzorów użytych do predykcji natężenia pola w terenach leśnych w modelu IŁ albo pozwolą na utworzenie nowych zależności empirycznych lub półempirycznych.

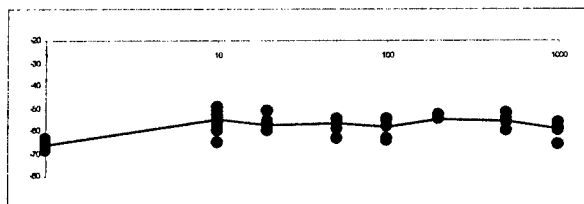
Po szczegółowej analizie lokalizacji terenów zalesionych w okolicach Wrocławia oraz usytuowania nadajników radiodifuzyjnych i radiokomunikacyjnych (GSM) dla potrzeb pomiarów wybrano dwa podobnej struktury obszary leśne, przy czym drugi z nich wykorzystywano tylko do pomiarów w paśmie 1800 MHz. Jako źródła sygnałów wybrane zostały nadajniki pracujące na częstotliwościach 104,9 MHz, 503,25 MHz, 936,5 MHz i 1836,5 MHz. Nadajnik telewizyjny (503,25 MHz) emitował sygnał o polaryzacji poziomej, pozostałe – o polaryzacji pionowej.

Pierwszą serię pomiarów prowadzono wczesną wiosną, w marcu 2003 r. Drzewa pozbawione były liści, o godz. 12.00 pierwszego dnia pomiarów temperatura wynosiła 0°C, drugiego i trzeciego dnia +2°C. Tylko te wyniki pomiarów prezentowane są w niniejszym referacie.

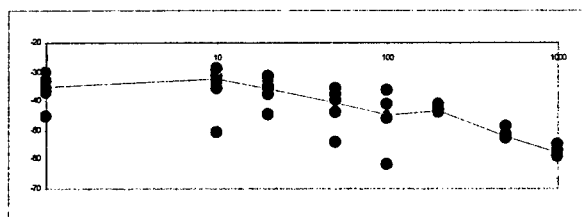
W obydwu badanych lasach wyznaczono trasy pomiarowe, na których dokonywane były pomiary natężenia pola elektrycznego selekcjonowane za pomocą anten kierunkowych od poszczególnych stacji nadawczych. Pomiary były przeprowadzane dla następujących odległości pomiarowych, licząc od skraju, w głąb lasu: 0 m – skraj lasu; 10 m; 20 m; 50 m; 100 m; 200 m; 500 m oraz 1000 m. Wstępnie planowano również pomiary dla odległości 2 km, ale w trakcie badań stwierdzono, że tłumienie sygnałów pochodzących od nadajnika GSM dla odległości przekraczających 1 km było w lesie tak duże, że poziom sygnału ginął w tle szumów.

Do badań użyto analizatora widma FSH3 Rohde&Schwarz nr ser. 100050 oraz zestawu anten aktywnych HE200 Rohde&Schwarz nr ser. 831096/030.

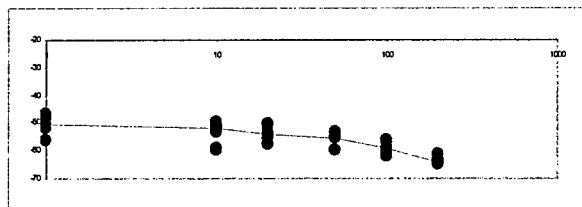
Wyniki odczytów zmian względnej mocy wybranych sygnałów radiowych w środowisku leśnym w funkcji odległości, dla poszczególnych pasm częstotliwości, zaprezentowane są na rysunkach 3 – 10.



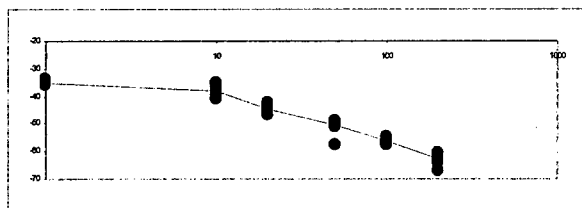
Rys. 3. Wyniki badań względnej mocy sygnału w paśmie 100 MHz, polaryzacja pionowa



Rys. 4. Wyniki badań względnej mocy sygnału w paśmie 500 MHz, polaryzacja pozioma



Rys. 5. Wyniki badań względnej mocy sygnału w paśmie 900 MHz, polaryzacja pionowa

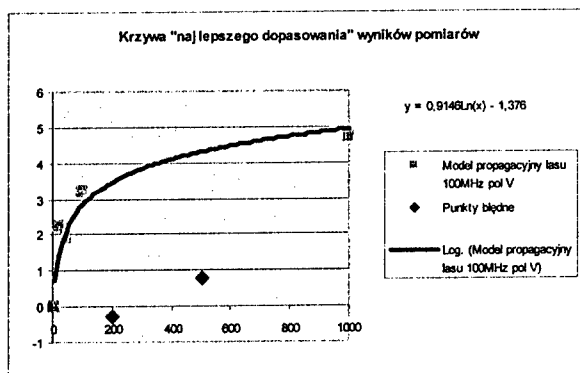


Rys. 6. Wyniki badań względnej mocy sygnału w paśmie 1800 MHz, polaryzacja pionowa

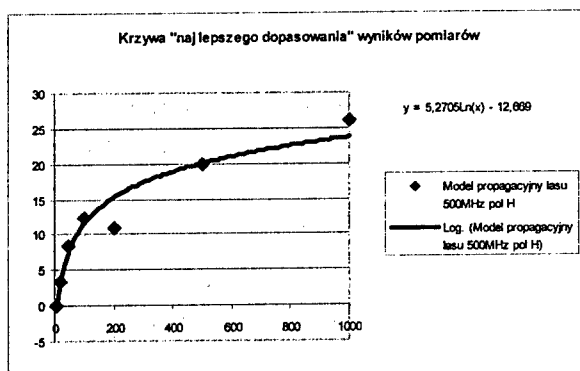
3. WNIOSKI

Analiza wyników pomiarów relacjonowanych w niniejszej pracy pozwala na stwierdzenie, że teoretyczny model uzyskany w wyniku prac prowadzonych w latach poprzednich praktycznie został potwierdzony. Potwierdzony został w szczególności charakter zmian tłumienia trasy w obszarach leśnych (przebieg tłumienia w funkcji odległości). Zaobserwowano jednak pewne anomalie w rozkładzie poziomu sygnału na odległościach bliskich skrajowi lasu. Potwierdza to obserwacje zrelacjonowane w dokumentach towarzyszących powstaniu Zalecenia 833, według których poziom ten stabilizuje się dopiero po przebyciu przez sygnał pewnej drogi wewnątrz lasu. Charakter zmian natężenia pola sygnału radiowego w terenie leśnym

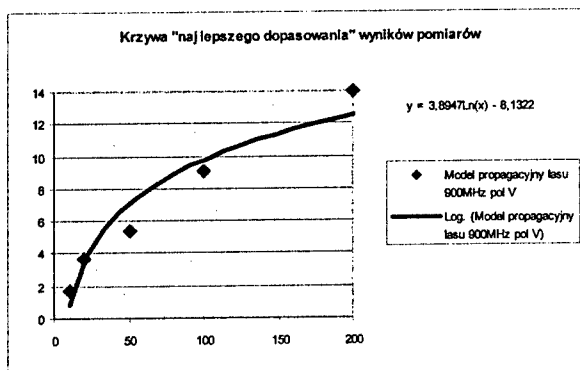
Krzywe „najlepszego dopasowania do wyników pomiarów” przedstawione są na rys. 7 – 10 różnią się swoim charakterem.



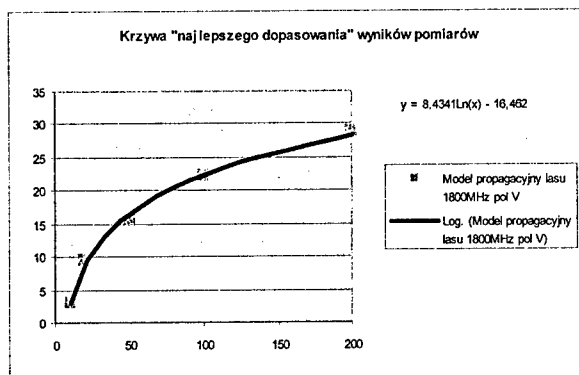
Rys. 7. Model propagacyjny lasu. Tłumienie sygnału w lesie, polaryzacja pionowa, $f=104,9$ MHz.



Rys. 8. Model propagacyjny lasu. Tłumienie sygnału w lesie, polaryzacja pozioma, $f=503,25$ MHz.



Rys. 9. Model propagacyjny lasu. Tłumienie sygnału w lesie, polaryzacja pionowa, $f=936,5$ MHz.



Rys. 10. Model propagacyjny lasu. Tłumienie sygnału w lesie, polaryzacja pionowa, $f=1863,5$ MHz.

Nadmienić należy, że w ramach pracy podjęto również interesujące, niepokazane tutaj, badania rozchodzenia się sygnału o polaryzacji poprzecznej do zamierzonej. Odbiór takiego sygnału będzie mieć miejsce, celowo lub przypadkowo, gdy płaszczyzna polaryzacji anteny odbiorczej usytuowana jest poprzecznie do płaszczyzny zamierzonej polaryzacji fali.

Wyniki relacjonowanych badań posłużyć mogą do bardziej precyzyjnego prognozowania zasięgów łączności radiowej w lasach, co istotne jest np. dla służb leśnych czy służb ratunkowych. Wyniki te mogą również być pomocne podczas koordynacji międzynarodowej stacji radiokomunikacji ruchomej lądowej, gdyż uwzględnienie dodatkowego tłumienia wnoszonego przez tereny zalesione spowoduje zmniejszenie liczby niezbędnych koordynacji. W szczególności, zweryfikowany model propagacji fal radiowych w terenach zalesionych włączony został do procedury prowadzonych w IŁ komputerowych obliczeń rozkładu natężenia pola i obliczeń zasięgowych systemów radiodifuzyjnych i radiokomunikacji ruchomej lądowej.

W referacie zawarto wybrane fragmenty ze sprawozdania z pracy nr 21 30 010 3 Instytutu Łączności [6]. Autor dziękuje firmie TESPOL Sp. z o. o. i Dolnośląskiemu Oddziałowi Okręgowemu URTiP we Wrocławiu za bezinteresowne udostępnienie analizatora widma i kompletu anten pomiarowych niezbędnych do wykonania pomiarów.

BIBLIOGRAFIA

- [1] ITU: Recommendation ITU-R P.833, *Attenuation in Vegetation*.
- [2] Tamir, T.: *Radio Wave Propagation along Mixed Paths in Forest Environments*. IEEE Trans. on AP, vol. AP-25, no. 4, July 1977.
- [3] Grzybowski M. J., *Model propagacji fali radiowej w terenach zalesionych*, KKRR'98, Poznań, str. 147 – 150.
- [4] ITU-R, Rec. P.526: *Propagation by diffraction*.
- [5] Grzybowski M. J., *Metoda oszacowania tłumienia fali radiowej wnoszonego przez lasy*, KKRRiT'2001, Poznań, str. 13.6-1 – 13.6.4.
- [6] Grzybowski M. J., Michalak M. P., Waga M. T., *Badanie tłumienności fali radiowej w obszarach leśnych*, Sprawozdanie Z21/21300103/841/2003 Instytut Łączności, Wrocław 2003.

LOSS OF RADIO WAVE PROPAGATION IN FORESTED ENVIRONMENT**Summary**

Scattering, diffraction and attenuation of radio wave in forested environment cause additional path loss. In ITU and IL models of propagation this loss is accounted by empirical corrections. These empirical corrections are defined by measurements. The ITU and IL models were tested. The radio wave loss measurement program was undertaken in mixed forests of Lower Silesia at frequencies from 100 MHz to 1800 MHz and at various distances varying from 10 m to 1000 m. The results of these studies are presented, discussed and utilize of the IL model to the wave propagation computer calculations supported by DTM is suggested.

Ryszard J. Katulski

Katedra Systemów i Sieci Radiokomunikacyjnych, Politechnika Gdańska

WSPÓŁCZESNA ŁĄCZNOŚĆ MORSKA

Streszczenie

W pracy opisano i scharakteryzowano systemy radiokomunikacyjne tworzące aktualny stan łączności morskiej w ramach globalnego systemu o nazwie GMDSS. Omówiono przeznaczenie i budowę tego systemu oraz scharakteryzowano jego elementy składowe, ze szczególnym uwzględnieniem morskiej łączności satelitarnej – w odniesieniu do systemów: INMARSAT oraz COSPAS-SARSAT. Poruszono zagadnienie rozgłaszania ostrzeżeń nawigacyjnych i meteorologicznych oraz opisano technikę radiowych środków ratunkowych w postaci radiopław awaryjnych i transponderów radarowych. Scharakteryzowano warunki propagacyjne występujące w łączności morskiej. Omówiono także zasady realizacji łączności podwodnej. W podsumowaniu wskazano trendy rozwojowe wynikające z ogólnego rozwoju techniki i technologii telekomunikacyjnej.

1. WPROWADZENIE

Systemowe rozwiązania łączności morskiej wynikają z kolejnych generacji ustaleń międzynarodowych, których nadrzędnym motywem jest podnoszenie stopnia bezpieczeństwa żeglugi, co jest uwarunkowane stopniem rozwoju techniki i technologii radiokomunikacyjnej. Aktualny stan łączności morskiej jest objęty wspólną nazwą: Globalny Morski System Łączności Alarmowej i Bezpieczeństwa, nazywany w skrócie **GMDSS** (*Global Maritime Distress and Safety System*). Z oczywistych względów ważnym elementem tego systemu jest łączność satelitarna, która ułatwia przekazywanie informacji na rozległych akwenach oceanicznych, co ma szczególne znaczenie w warunkach zagrożenia [1, 2].

2. SYSTEM GMDSS

2.1. Budowa i charakterystyka użytkowa

System GMDSS obok tego, że zapewnia tzw. łączność ogólną związaną z eksploatacją statków, jest w szczególności przeznaczony do realizacji poniżej wymienionych zadań:

- alarmowanie o zagrożeniu,
- lokalizowanie miejsc katastrof morskich,
- wspomaganie akcji ratunkowych,
- rozpowszechnianie informacji związanych z bezpieczeństwem żeglugi.

Do telekomunikacyjnej obsługi tych zadań służą następujące systemy radiokomunikacyjne:

- geostacjonarny system satelitarny **INMARSAT** (*International Maritime Satellite Communications*),
- biegunowy system satelitarny **COSPAS-SARSAT** (*Search and Rescue Satellite-Aided Tracking System*),
- satelitarne radiopławy awaryjne **EPIRB** (*Emergency Positioning Indicating Radio Beacon*),
- satelitarny system wywołania grupowego **EGC** (*Enhanced Group Calling*),
- cyfrowe selektywne wywołanie **DSC** (*Digital Selective Calling*),
- systemy rozgłaszania ostrzeżeń – **NAVTEX** (*Navigational Telex*) i **Safety NET**, oraz oczywiście także
- naziemna radiotelefonii.

Przepisy klasyfikacyjne określające wymagania związane z budową i wyposażeniem statków, dotyczą także ich obowiązkowego wyposażenia telekomunikacyjnego, które jest związane z zasięgiem pływania każdego statku. Z tego względu przyjęto poniższy podział akwenów pływania:

- rejon morski A1, który wynika z zasięgu działania stacji brzegowej ultrakrótkofalowej – 156.525 MHz, co obejmuje obszar przybrzeżny w odległości do 30 Mm (mil morskich) od linii brzegowej,
- rejon morski A2, który wynika z zasięgu działania stacji brzegowej pośredniofalowej – 2187.5 kHz, z wyłączeniem rejonu A1, co obejmuje obszar od 30 do 150 Mm od brzegu,
- rejon morski A3, który wynika z zasięgu działania systemu INMARSAT, z wyłączeniem rejonów A1 i A2, co odnosi się do żeglugi dalekomorskiej, nie obejmującej jednak obszarów podbiegunowych,
- morski rejon A4, który obejmuje obszary podbiegunowe, które nie są objęte zasięgiem działania wymienionego satelitarnego systemu geostacjonarnego.

Wynikający z powyższego standard wyposażenia statków w urządzenia telekomunikacyjne przedstawia się następująco:

- statki przeznaczone do żeglugi tylko w rejonie A1 muszą być wyposażone w radiostację UKF do obsługi kanałów radiowych w paśmie od 156 do 162.05 MHz, przeznaczonych do komunikacji ogólnej i publicznej, która w szczególności powinna obsługiwać wydzielone kanały specjalne przeznaczone do pracy w niebezpieczeństwie i dla potrzeb bezpieczeństwa, tzn.:
 - kanał nasłuchowy z zastosowaniem DSC na częstotliwości 156.525 MHz, przeznaczony do odbioru sygnałów alarmowych,
 - kanał radiotelefoniczny na częstotliwości 156.8 MHz, przeznaczony do wspomagania akcji ratunkowych,
- statki przeznaczone do pływania najdalej w rejonie A2, obok wyposażenia właściwego dla rejonu A1, muszą być wyposażone w radiostację pośredniofalową do obsługi kanałów radiowych w paśmie od 1605 do 4000 kHz, która w szczególności powinna obsługiwać wydzielone kanały specjalne, tzn.:
 - kanał radioteleksowy na częstotliwości 2174.5 kHz,
 - kanał radiotelefoniczny na częstotliwości 2182 kHz,
 - kanał nasłuchowy z zastosowaniem DSC na częstotliwości 2187.5 kHz, przeznaczony do odbioru sygnałów alarmowych,
 - kanał radiotelefoniczny na częstotliwości 3023 kHz, przeznaczony do komunikacji z samolotem biorącym udział w akcji ratunkowej,

- statki dalekomorskie, tzn. przeznaczone do pływania w strefie A3, obok wyposażenia właściwego dla rejonów A1 i A2, muszą być wyposażone w radiostację do łączności satelitarnej w paśmie 1.5/1.6 GHz – terminal INMARSAT-u,
- natomiast statki przeznaczone do pływania w strefie A4, obok powyższego wymienionego zestawu urządzeń, powinny być dodatkowo wyposażone w radiostację krótkofalową.

Przedstawiony powyżej zakres wyposażenia statków w urządzenia telekomunikacyjne nie wyczerpuje w pełni obowiązujących wymogów klasyfikacyjnych, o czym będzie mowa w dalszych częściach artykułu.

2.2. Morska łączność satelitarna

Aktualnie ogólnie dostępna morska łączność satelitarna opiera się głównie na systemie INMARSAT, który zapewnia dostęp do usług telekomunikacyjnych wszelkim użytkownikom, w tym także morskim, rozrzuconym na rozległym obszarze pomiędzy szerokościami geograficznymi +/- 75 stopni [3].

Podstawą działania tego systemu są cztery spośród dziewięciu telekomunikacyjnych satelitów geostacjonarnych o dwóch standardach, tzn.:

- satelity INMARSAT-2 o przepustowości 250 kanałów fonicznych, z wiązkami globalnymi, oraz
- satelity INMARSAT-3 o przepustowości 2000 kanałów fonicznych, z wiązkami globalnymi i regionalnymi do obsługi tanich terminali osobistych.

Rolę nadrzędną w działaniu tego systemu, która polega na nadzorowaniu i koordynowaniu pracy satelitów oraz stacji naziemnych i terminali użytkowników, spełnia Centrum Kontroli Sieci **NCC** (*Network Control Centre*) znajdujące się w Londynie. Natomiast bezpośrednia kontrola satelitów na orbicie odbywa się poprzez trzy Centra Kontroli Satelitów **SCC** (*Satellite Control Centre*), które zostały zainstalowane w Waszyngtonie, Darmstadt i Londynie.

Z punktu widzenia potrzeb telekomunikacyjnych w praktyce pływania na morzu, system INMARSAT pozwala realizować następujące usługi telekomunikacyjne:

- przede wszystkim łączność w niebezpieczeństwie, we współpracy z właściwym dla danego akwenu Ratowniczym Centrum Koordynacyjnym systemu GMDSS,
- wywołanie grupowe do statków przebywających w wybranym akwenu morskim,
- radiotelefonie i radiotelegrafię dalekopisową, oraz
- dwukierunkową transmisję danych.

Wymienione powyżej usługi są wykonywane przy użyciu terminali o bardzo zróżnicowanych standardach, co przedstawia się następująco:

- analogowy standard INMARSAT-A do transmisji mowy i modemowej transmisji danych w kanałach fonicznych oraz do usług teleksowych i telefaksowych – zostanie wyłączony przed końcem 2007 roku,
- standard INMARSAT-A HSD do szybkiej jednokierunkowej transmisji danych do stacji brzegowej, z szybkością 64 kbit/s,
- standardy INMARSAT-B oraz INMARSAT-B HSD, będące cyfrowymi odmianami wymienionych powyżej standardów analogowych,
- standard INMARSAT-C do transmisji danych tekstowych,
- małogabarytowe terminale nadawczo-odbiorcze lub tylko odbiorcze (paging) w standardach INMARSAT-D/D+, wymagające użycia na satelicie wiązek regionalnych,
- satelitarne radiopławy awaryjne EPIRB w standardzie INMARSAT-E, oraz

- standardy INMARSAT-M/mini M do cyfrowej transmisji sygnałów mowy, przy czym „mini M” wymaga także użycia wiązek regionalnych.

Strategia rozwoju INMARSAT-u zmierza do zapewnienia użytkownikom tego systemu nowych wysokiej jakości usług telekomunikacyjnych. Przykładem tego jest wprowadzony do eksploatacji system o nazwie INMARSAT-GAN (*Global Area Network*), który oferuje użytkownikom – także ruchomym – dostęp poprzez satelitę do:

- stacjonarnej cyfrowej sieci z integracją usług typu ISDN – usługa Mobile ISDN,
- sieci INTERNET – usługa Mobile IP.

W najbliższej przyszłości planuje się ewolucję systemu GAN w kierunku **B-GAN** (*Broadband GAN*), co oznacza realizację multimedialnych usług szerokopasmowych **PMC** (*Personal Multimedia Communications*).

Obok typowego systemu komunikacyjnego jakim jest INMARSAT, w łączności morskiej stosowany jest także wymieniony wcześniej system COSPAS-SARSAT, przeznaczony do alarmowania i lokalizacji statków znajdujących się w niebezpieczeństwie. W przeciwieństwie do INMARSAT-u, działanie tego systemu opiera się o satelity umieszczone na orbitach biegunowych, z czego wynika jego globalny zasięg działania – obejmujący także obszary podbiegunowe. System ten tworzą w ramach GMDSS-u dwa niezależne systemy (z czego pochodzi jego nazwa), tzn.:

- kosmiczny system do poszukiwania statków (COSPAS) z orbitami na wysokości 1000 km, zbudowany przez byłych Związek Radziecki, oraz
- satelitarny system do poszukiwania i ratownictwa (SARSAT) z orbitami na wysokości 850 km, zbudowany wspólnie przez Stany Zjednoczone, Kanadę i Francję.

W obu tych systemach stosowane są dwa podpasma częstotliwości: 121.5 oraz 406 MHz. W strukturze GMDSS-u oba systemy służą do lokalizowania radiopław awaryjnych EPIRB, nadających sygnały alarmowe. Po odebraniu tych sygnałów przez urządzenia na satelicie, są one przekazywane do Ratowniczego Centrum Koordynacyjnego, za pośrednictwem lokalnej stacji naziemnej.

2.3. Rozgłaszanie ostrzeżeń nawigacyjnych i meteorologicznych

W celu podniesienia stopnia bezpieczeństwa żeglugi, do eksploatacji wprowadzono dwa systemy radiokomunikacyjne przeznaczone do rozgłaszania informacji związanych z bezpieczeństwem żeglugi, które dotyczą wszelkich ostrzeżeń nawigacyjnych lub meteorologicznych oraz krótko i długoterminowych prognoz pogodowych.

Pierwszy z wymienionych - system NAVTEX – pracuje w paśmie fal średnich, na częstotliwości 518 kHz. Odbiór informacji w tym systemie odbywa się przy użyciu specjalnych odbiorników radiotelegraficznych, w które muszą być wyposażone wszystkie statki o pojemności powyżej 300 BRT (niem. *Bruttoregistertonne*), niezależnie od rejonu pływania. Odbiór komunikatów bezpieczeństwa odbywa się automatycznie, łącznie z ich wydrukiem. Pozwala to na bezobsługowy całodobowy nasłuch.

Drugi z wymienionych – system Safety NET – wchodzi w skład satelitarnego systemu wywołania grupowego EGC, który poprzez INMARSAT umożliwia autoryzowanym instytucjom przysyłanie komunikatów do wybranej grupy statków. Rozróżnia się cztery poziomy ważności tych komunikatów:

- niebezpieczeństwo (distress) – komunikat alarmowy,
- pilne (urgent),
- informacja bezpieczeństwa (safety) – komunikat ostrzegawczy,
- informacja zwykła (routine).

2.4. Radiowe środki ratunkowe

Stosowane obowiązkowo na statkach, niezależnie od rejonu pływania, radiowe środki ratunkowe to: radiopławy awaryjne oraz transpondery radarowe.

System satelitarnych radiopław awaryjnych EPIRB jest przeznaczony do bezwłocznego powiadamiania o wystąpieniu sytuacji zagrożenia. Radiopławy te pracują w następujących zakresach częstotliwościowych:

- w paśmie od 406 do 406.1 MHz, we współpracy z systemem COSPAS/SARSAT, oraz
- w paśmie od 1.6453 do 1.6455 GHz, we współpracy z systemem INMARSAT.

Zadaniem radiopławy jest wysyłanie sygnału radiowego zawierającego następujące informacje:

- aktualną pozycję zagrożonego statku,
- jego dane identyfikacyjne,
- inne dodatkowe informacje ułatwiające prowadzenie akcji ratunkowej.

Każda radiopława jest tak skonstruowana, że może być uruchomiona ręcznie na statku lub automatycznie po spłynięciu na wodę. Po takim uaktywnieniu nadajnik radiopławy wysyła w kierunku satelity sygnał radiowy o mocy 1 W, poprzez antenę o hemisferycznej charakterystyce promieniowania. Ponadto, po uruchomieniu nadajnika uaktywnia się również zainstalowany na radiopławie transponder radarowy, pracujący jako urządzenie odziewowe.

Przeznaczeniem transpondera radarowego jest pomoc w wykrywaniu i lokalizowaniu rozbitków na morzu. Odbywa się to przy użyciu specjalnego nadajnika radiowego, który uaktywnia się pod wpływem docierających do niego sygnałów radarowych wysyłanych z ze statków lub samolotów. Na każdy sondujący impuls radarowy nadajnik transpondera wysyła odpowiedź w postaci krótkotrwałego sygnału o długości 100 μ s, z szybkim przesuwaniem częstotliwości w paśmie od 9.2 do 9.5 GHz. Taki odebrany sygnał odzewowy jest zobrazony na ekranie radaru w postaci równo oddalonych kropek, tworząc tzw. kreskę namiarową wskazującą położenie transpondera.

3. UWARUNKOWANIA PROPAGACYJNE W ŁĄCZNOŚCI MORSKIEJ

3.1. Rejon morski A1 – propagacja fal ultrakrótkich

W systemach łączności morskiej przybrzeżnej jest stosowany zakres fal ultrakrótkich – co napisano wcześniej, rozchodzących się w postaci fali przestrzennej. Tego rodzaju mechanizm propagacyjny ma miejsce wówczas, gdy przynajmniej jedna z anten łącza radiowego jest wzniesiona na wysokość równą co najmniej dziesięciu długościom fali radiowej. W omawianej sytuacji kryterium to zawsze spełnia zawieszenie anteny stacji brzegowej. W tym miejscu należy podkreślić, że w celu spełnienia wymagań zasięgowych, w omawianej sytuacji konieczne jest nie tylko zapewnienie określonego poziomu mocy sygnału radiowego lecz także umieszczenie anteny stacji brzegowej na odpowiedniej wysokości, gwarantującej spełnienie tych wymagań – uzyskując w praktyce zasięg quasi-horyzontowy.

W porównaniu do warunków lądowych, analiza propagacyjna fali przestrzennej w radiowym łączu morskim z jednej strony jest uproszczona z uwagi na to, że można w niej pominąć wpływ ukształtowania podłoża trasy propagacji oraz infrastruktury środowiska propagacyjnego. Jednakże z drugiej strony należy wziąć pod uwagę zjawisko odbicia fali

radiowej od powierzchni morza, wskutek czego do anteny na statku docierają dwa rodzaje fal, tzn. fala bezpośrednia oraz zazwyczaj kilka fal odbitych, będących skutkiem propagacji wielodrogowej. W rezultacie tego mogą mieć miejsce zaniki odbieranego sygnału, co prowadzi do obniżenia jakości transmisji radiowej. Najmniej korzystna sytuacja ma miejsce w warunkach gładkiej powierzchni spokojnego morza. Natomiast najlepiej pod tym względem jest wówczas, gdy morze jest wzburzone.

Podstawowym modelem stosowanym przy projektowaniu ultrakrótkofalowych systemów łączności morskiej są statystyczne krzywe propagacji tych fal zalecane przez ITU-R. Wyrażają one rozkład wartości natężenia pola elektrycznego w funkcji odległości, przy czym wartości te są związane z określonym prawdopodobieństwem przestrzennym i czasowym ich wystąpienia w praktyce.

3.2. Rejon morski A2 – propagacja fal pośrednich

W systemach łączności morskiej w rejonie A2 jest stosowany zakres pośredniofalowy, w którym sygnał radiowy rozchodzi się w postaci fali powierzchniowej, na której propagację w dominującym stopniu mają wpływ parametry elektryczne podłoża – w tym przypadku wody morskiej. Tego rodzaju mechanizm propagacyjny ma miejsce wówczas, gdy przynajmniej jedna z anten łącza radiowego jest w sensie elektrycznym położona blisko podłoża trasy propagacji, a więc wysokość wzniesienia tej anteny jest dużo mniejsza od dziesięciu długości fali radiowej. W omawianej sytuacji kryterium to spełniają zawieszenia obu anten łącza radiowego.

Parametry elektryczne wody morskiej, które determinują wielkość natężenia pola fali powierzchniowej, to przenikalność elektryczna oraz przewodność właściwa, których wartości są dużo większe aniżeli ma to miejsce w przypadku trasy lądowej. Wpływa to korzystnie na propagację fali radiowej nad wodą.

Ponadto, należy podkreślić, że w omawianej sytuacji, obok obszaru pokrycia wynikającego z zasięgu oddziaływania fali powierzchniowej, może wystąpić niekorzystne zjawisko wynikające z odbicia od jonosfery fali wypromieniowanej przez wiązki boczne anteny.

3.2. Rejon morski A3 – propagacja mikrofal w łączu satelitarnym

Trasa propagacji fali radiowej w łączu satelitarnym jest nieporównywalnie dłuższa od tego co ma miejsce w łączu naziemnym. Zatem w takiej sytuacji, istotnym składnikiem tłumienia propagacyjnego tej fali jest zjawisko jej rozproszenia, co w przypadku łączy radiowych w systemie INMARSAT może dochodzić nawet do 190 dB.

Ponadto mając na uwadze fakt, że fala radiowa w łączu satelitarnym przechodzi przez troposferę i jonosferę, należy łącznie uwzględnić następujące zjawiska:

- tłumienie propagacyjne w wolnej przestrzeni, wynikające z długości trasy propagacji oraz częstotliwości sygnału radiowego,
- refrakcję i absorpcję troposferyczną,
- oddziaływanie jonosfery, oraz
- odbicia od powierzchni morza.

Skutkiem występowania refrakcji troposferycznej jest zmiana trajektorii fali radiowej, co jest zależne od parametrów radiometeorologicznych, tzn. od ciśnienia powietrza i jego temperatury. W wyniku tego obserwuje się tzw. błąd refrakcji powodujący pozorne zwiększenie tzw. wysokości katowej położenia satelity. Natomiast absorpcja troposferycz-

na ma swoje źródło w opadach i zawiesinach wodnych w postaci chmur, z czego wynika dodatkowe osłabienie fali radiowej [4].

3.2. Rejon morski A4 – propagacja fal krótkich

Łączność krótkofalowa z natury rzeczy jest stosowana przede wszystkim do realizacji komunikacji dalekosiężnej, wykorzystując do tego zjawisko odbicia się fali radiowej od jonosfery. Ten rodzaj mechanizmu propagacyjnego charakteryzuje się zmianami dobowymi i rocznymi, co czyni ważnym stosowny do tego dobór częstotliwości sygnału radiowego, tzn: optymalnej częstotliwości roboczej **FOT** (*Frequency Optimum Traffic*), która powinna być położona pomiędzy najmniejszą częstotliwością użytkową **LUF** (*Lowest Usable Frequency*) i maksymalną częstotliwością użytkową **MUF** (*Maximum Usable Frequency*).

4. ŁĄCZNOŚĆ PODWODNA

Zainteresowanie łącznością podwodną wynika głównie z potrzeb wojskowych, przede wszystkim do komunikowania się z okrętami podwodnymi, przy czym rozróżnia się [5]:

- jednokierunkową łączność strategiczną na bardzo duże odległości, przy użyciu bardzo długich fal radiowych, oraz
- dwukierunkową łączność taktyczną na krótkie odległości, przy użyciu fal akustycznych.

Powyższe rodzaje łączności podwodnej są aktualnie stosowane w praktyce morskiej. Jednakże od wielu lat prowadzone są zaawansowane prace zmierzające do zastosowania w podwodnej łączności strategicznej fal optycznych, wykorzystując do tego pośrednictwo segmentu satelitarnego [6].

Fale radiowe z natury rzeczy źle wnikają w środowisko wodne, dlatego łączność przy ich użyciu może być realizowana na względnie niewielką głębokość – rzędu kilkaset metrów, stosując do tego możliwie jak najmniejszą częstotliwość. W takiej sytuacji, jednokierunkowe łącze radiowe od naziemnego centrum dowodzenia do zanurzonego okrętu podwodnego, znajdującego się odległości co najmniej kilku tysięcy kilometrów od tego centrum, składa się z dwóch części:

- bardzo długiego odcinka nawodnego, oraz
- stosunkowo krótkiego odcinka w wodzie morskiej.

Zadaniem odcinka nawodnego jest wytworzenie przy powierzchni morza, w miejscu zanurzenia okrętu podwodnego, pola elektromagnetycznego o odpowiedniej wartości, umożliwiającej wniknięcie tego pola na głębokość zanurzenia okrętu. Aby zapewnić efektywność takiego procesu transmisji, nadawany sygnał musi mieć odpowiedni – bardzo duży, rzędu MW – poziom energetyczny.

Stosowanie fal radiowych o bardzo małych częstotliwościach wymaga użycia na okręcie podwodnym odpowiedniego urządzenia antenowego, którego zadaniem jest spełnianie roli detektora składowej elektrycznej pola elektromagnetycznego. Antena taka najczęściej ma postać rozciągniętego w wodzie przewodu, w którym indukuje się napięcie proporcjonalne do wielkości dochodzącego do tego przewodu pola elektrycznego.

Przenoszona w takich warunkach informacja ma postać umownych sekwencji znaków binarnych przesyłanych z bardzo małą prędkością.

Natomiast podwodna łączność taktyczna polega na dwukierunkowej transmisji informacji pomiędzy operującymi w zespole okrętami podwodnymi i nawodnymi. Zadanie to

realizuje się przy użyciu tzw. hydrotelefonu, którego działanie oparte jest na falach akustycznych, w paśmie od kilku do kilkudziesięciu kHz.

Zastosowanie w łączności podwodnej kanału optycznego jako alternatywy kanału radiowego, wynika przede wszystkim z potrzeby podwyższenia dotychczasowej szybkości transmisji. Wyniki uzyskane w trakcie wykonanych dotąd badań potwierdziły wyższość pod tym względem kanału optycznego.

5. ZAKOŃCZENIE

Z eksploatacyjnego punktu widzenia, jak to zaznaczono na wstępie, ważnym elementem w systemie łączności morskiej jest system INMARSAT, który podlega ciągłemu rozwojowi i rozbudowie. Docelowo ma on sprostać spodziewanemu zapotrzebowaniu na różne rodzaje zaawansowanych usług telekomunikacyjnych, z konieczną do tego szybkością cyfrową transmisją szerokopasmową.

Do realizacji tych zamierzeń planuje się umieścić na orbicie kolejną generację satelitów klasy INMARSAT-4, z kanałami szerokopasmowymi o przepustowości do 432 kbit/s. Zakłada się przy tym, że zastosowane wówczas protokoły komunikacyjne będą kompatybilne ze standardem trzeciej generacji.

Reasumując, przyszły stan łączności morskiej, w szczególności do obsługi żeglugi dalekomorskiej na otwartych akwenach oceanicznych, będzie wynikał z aktualnego rozwoju telekomunikacyjnej technologii satelitarnej [7].

BIBLIOGRAFIA

- [1] Katulski R.J., Kosiński A., Pawłowski W., Żurek J.: *Stan obecny i perspektywy radiokomunikacji morskiej*, Przegląd Telekomunikacyjny, 12/1999, 816-820.
- [2] Katulski R.J., Kosiński A., Żurek J.: *Perspektywy rozwoju nowoczesnej radiokomunikacji morskiej*, Przegląd Telekomunikacyjny, 11/2001, 745-749.
- [3] Katulski R.J.: *Rozwój i możliwości użytkowe systemu INMARSAT*, Przegląd Telekomunikacyjny, 12/2000, 913-917.
- [4] Katulski R.J.: *Łączność stratosferyczna*, Przegląd Telekomunikacyjny, 5/2003, 249-253.
- [5] Katulski R.J., Kosiński A., Pawłowski W., Studański R.: *Podstawy teorii i techniki łączności podwodnej*, Przegląd Telekomunikacyjny, 9/1998, 621-629.
- [6] Katulski R.J., Kosiński A.: *Podwodna łączność optyczna*, Przegląd Telekomunikacyjny, 2-3/2004, 99-102.
- [7] Katulski R.J., Mikołajski M.: *Uwarunkowania rozwojowe w telekomunikacji satelitarnej*, Przegląd Telekomunikacyjny, 7/2003, 370-375.

CONTEMPORARY MARITIME COMMUNICATIONS

Summary

In the paper the maritime communication system called GMDSS is described. The functional assignment of the system is characterized. The satellite communication systems, i.e. INMARSAT and COSPAS-SARSAT, are presented. The problem of underwater communication is also described. In conclusion, the development trends are outlined.

Ryszard J. Katulski*, Andrzej Kiedrowski**

**Katedra Systemów i Sieci Radiokomunikacyjnych*,
Netia Telekom S.A.****

ZMODYFIKOWANY MODEL COST 231 HATY DO PROJEKTOWANIA SYSTEMÓW DOSTĘPOWYCH

Streszczenie

W pracy przedstawiono zagadnienie wyznaczania strat propagacyjnych sygnału radiowego w radiowym łączu dostępowym pracującym w warunkach miejskich. Podano zmodyfikowaną postać modelu COST 231 HATY, opracowaną na podstawie wykonanych obszernych badań eksperymentalno-pomiarowych, przeprowadzonych w dużych aglomeracjach miejskich na terenie całego kraju. Podano nowe podejście przy określaniu parametrów wejściowych tego modelu, wynikające z geometrii tras propagacji sygnału radiowego, położonych ponad zabudową podłoża propagacyjnego. W podsumowaniu określono zakres stosowności tego modelu..

1. WPROWADZENIE

Koncepcja radiowego systemu dostępowego polega na zastosowaniu specjalnej stacji radiokomunikacyjnej, nazywanej stacją bazową lub główną, poprzez którą użytkownik najczęściej nieruchomy uzyskuje bezprzewodowy dostęp do usług oferowanych przez stacjonarną sieć telekomunikacyjną. Ma to szczególne znaczenie wówczas, gdy niemożliwe jest zadawalająco szybkie zrealizowanie takiego dostępu w sposób przewodowy.

Ważnym elementem procedury projektowania radiowego systemu dostępowego jest określenie strat propagacyjnych sygnału radiowego w łączu dostępowym, pomiędzy stacją bazową i stacją abonencką, będącą na wyposażeniu użytkownika. Jest to szczególnie kłopotliwe w warunkach pracy tego systemu w terenie zabudowanym, co najczęściej ma miejsce w praktyce.

Istnieje wiele modeli przeznaczonych do szacowania tłumienia propagacyjnego w warunkach miejskich. Modele te zostały wykonane dla potrzeb radiokomunikacji ruchomej lądowej (RLL), kiedy to z jednej strony antena stacji bazowej jest usytuowana ponad dachami budynków, natomiast z drugiej strony antena terminala użytkownika znajduje się pomiędzy budynkami na poziomie ulicy. Z tych powodów modele te nie uwzględniają specyfiki propagacyjnej występującej w stałym łączu dostępowym, w którym obie anteny tego łącza są położone ponad budynkami.

Wyniki przeprowadzonych dotychczas badań eksperymentalno-pomiarowych, obejmujące najbardziej reprezentatywne z tych modeli, pokazują, że w warunkach pracy

radiowego łącza dostępowego najbardziej użytecznym z nich wydaje się być model COST 231 Walfisha-Ikegami [1].

Jednakże opisany w tym modelu sposób wyznaczania strat propagacyjnych charakteryzuje się dużą złożonością, szczególnie w sytuacjach przysłonięcia linii bezpośredniej widoczności anten, co wymaga wykonania bardzo dokładnej analizy wpływu geometrii zabudowy położonej na trasie propagacji sygnału radiowego. Dużo łatwiejszy w użyciu jest inny z przebadanych modeli, tzn. model COST 231 Haty, który jednak dużo gorzej przystaje do warunków występujących w łączu dostępowym. W takim stanie rzeczy uzasadnione jest dążenie do modyfikacji tego modelu, pozwalającej na jego zastosowanie do projektowania takiego łącza [2].

W pracy przedstawiono nową postać modelu COST 231 HATY, uzyskaną w wyniku obszernych badań przeprowadzonych na podstawie przeszło ośmiu tysięcy pomiarów, wykonanych w dużych miastach na terenie całego kraju. Podano nowe podejście przy określaniu parametrów wejściowych tego modelu, wynikające z geometrii tras propagacji sygnału radiowego, położonych ponad zabudową podłoża propagacyjnego. Określono stosowny do tego zakres stosowalności tego modelu.

2. MODEL COST 231 HATY

2.1. Forma klasyczna dla systemów RLL

Europejska grupa badawcza powołana w latach 90-tych w ramach programu COST (Cooperation in Science and Technology), w części zajmującej się zagadnieniami rozchodzenia się fal radiowych w systemach radiokomunikacji ruchomej w terenie zabudowanym, opracowała przystającą do europejskich warunków propagacyjnych zmodyfikowaną postać modelu Okumury-Haty, o nazwie: model COST 231 Haty. Według tego modelu tłumienie propagacyjne L w ruchomym łączu radiowym jest określone poniższą zależnością:

$$L[\text{dB}] = 46,3 + 33,9\lg(f[\text{MHz}]) - 13,82\lg(h_B[m]) - a(h_T) + \\ + (44,9 - 6,55\lg(h_B[m]))\lg(d[\text{km}]) + C, \quad (1)$$

przy czym:

$$a(h_T) = (1,1\lg(f[\text{MHz}]) - 0,7)h_T[m] - (1,56\lg(f[\text{MHz}]) - 0,8), \quad (1.1)$$

gdzie h_B i h_T oznaczają wysokości położenia anten, odpowiednio stacji bazowej i terminala użytkownika. Natomiast stała C przyjmuje jedną z dwóch wartości:

- dla terenów podmiejskich i średnich miast $C = 0 \text{ dB}$,
- dla centrów dużych miast $C = 3 \text{ dB}$.

Model ten został opracowany na podstawie pomiarów wykonanych w warunkach europejskiej zabudowy miejskiej, przy czym warunki techniczne tych pomiarów były następujące:

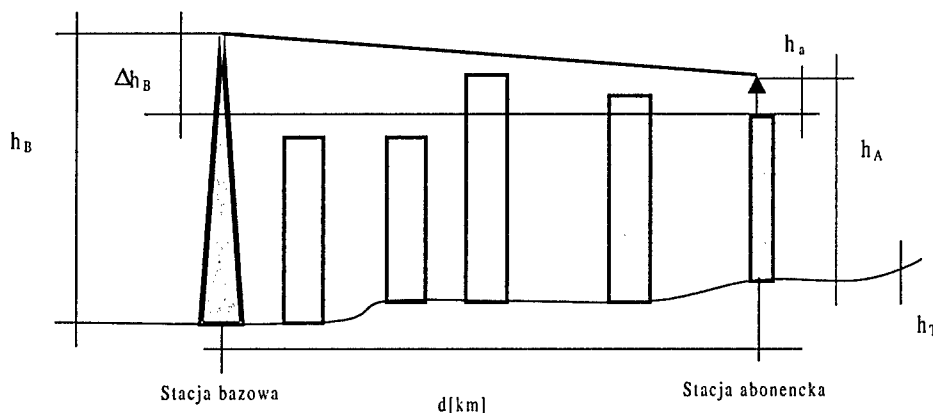
- długości tras propagacyjnych d zawierały się w granicach od 1 do 20 km,
- zakres częstotliwości f , który pierwotnie wynosił od 150 do 1000 MHz, został rozszerzony do 2 GHz,

- wysokości h_B zawieszenia anten stacji bazowych, określone względem powierzchni podłoża propagacyjnego, zawierały się w zakresie od 30 do 200 m npt (nad poziomem terenu),
- natomiast wysokości h_T położenia anten terminali użytkowników, określone względem poziomu ulicy, wynosiły od 1 do 10 m npt.

Powyższa charakterystyka techniczna określa zakres stosowalności klasycznej formy opisywanego modelu.

2.2. Geometria tras propagacyjnych w systemach dostępowych

Typową geometrię tras propagacyjnych w radiowych systemach dostępowych, pracujących w warunkach miejskich, przedstawiono na rys. 1. Na rysunku tym pokazano przekrój środowiska propagacyjnego pomiędzy stacją bazową i abonencką, uwzględniając przy tym znajdujące się tam elementy charakterystycznej zabudowy miejskiej.



Rys. 1. Typowa geometria tras propagacyjnych w radiowych systemach dostępowych

Na powyżej zilustrowanej trasie propagacyjnej zaznaczono charakteryzujące ją parametry, które są następujące:

- długość d tej trasy pomiędzy antenami obu stacji,
- wysokości h_B oraz h_A zawieszenia npt anten, odpowiednio stacji bazowej i abonenckiej,
- znormalizowaną wysokość Δh_B zawieszenia anteny stacji bazowej,
- wysokość h_a zawieszenia anteny stacji abonenckiej nad poziomem dachu,
- wysokość h_T podłoża propagacyjnego, określoną względem powierzchni morza (npm).

Jak widać, w warunkach pracy systemu dostępowego w mieście, obie anteny łączy radiowego są położone ponad dachami budynków, przy czym anteny stacji abonenckich instalowane są na niewielkiej wysokości nad powierzchnią dachu – zazwyczaj około 3 m.

- Zatem w analizowanej sytuacji, z punktu widzenia działania systemu dostępowego, istotna jest ta część fali radiowej, która rozchodzi się ponad dachami budynków. Jest to sytuacja odmienna od tej z jaką mamy do czynienia w przy-

padku systemu RRL, kiedy to istotna jest ta część fali radiowej, która wnika pomiędzy budynki i dociera do terminala użytkownika poruszającego się pomiędzy budynkami.

Należy w tym miejscu podkreślić, że sytuacja propagacyjna pokazana na rysunku 1, ilustruje stan bezpośredniej widoczności anten obu stacji łączy radiowego, tzn. stan LOS (Line-of-Sight), co zazwyczaj ma miejsce w praktyce budowy radiowych systemów dostępowych. Nie wyklucza to jednak innych sytuacji, kiedy może mieć miejsce przysłonięcie linii łączącej obie anteny, tzn. stan NLOS (Non Line-of-Sight).

2.3. Forma zmodyfikowana dla systemów dostępowych

Mając na uwadze powyższe uwarunkowania, proponuje się:

- po pierwsze, zmianę w sposobie określania parametrów technicznych łączy radiowego, występujących w klasycznej postaci modelu COST 231 Haty, oraz
- po drugie, zmodyfikowaną postać analityczną tego modelu, w porównaniu do tej jak to opisują zależności (1 i 1.1).

Podstawą tych zmian jest inne określenie wysokości zawieszenia anteny stacji bazowej, znormalizowanej względem poziomu dachu, na którym jest zainstalowana antena stacji abonenckiej, co wyraża się następująco – patrz rys. 1:

$$\Delta h_B = h_B - (h_A - h_a). \quad (2)$$

Konsekwentnie do tego, wysokość zawieszenia anteny stacji abonenckiej należy określać także względem tego poziomu, tzn. jako h_a .

Zmodyfikowaną postać analityczną modelu COST 231 Haty opracowano poprzez analizę regresji liniowej funkcji błędu pomiędzy pomierzonymi i obliczonymi wartościami tłumienia propagacyjnego, o czym będzie mowa w następnym rozdziale.

W rezultacie otrzymano nową formę modelu COST 231 Haty, która jest określona następującą zależnością:

$$\begin{aligned} L[\text{dB}] = & 46,3 + 33,9 \lg(f[\text{MHz}]) - 13,82 \lg(\Delta h_B [m]) - a(h_a) + \\ & + (44,9 - 6,55 \lg(\Delta h_B [m])) \lg(d[\text{km}]) + \\ & + (-27,54 - 14,85 \lg(d[\text{km}]) + 16,04 \lg((\Delta h_B [m])/30)), \end{aligned} \quad (3)$$

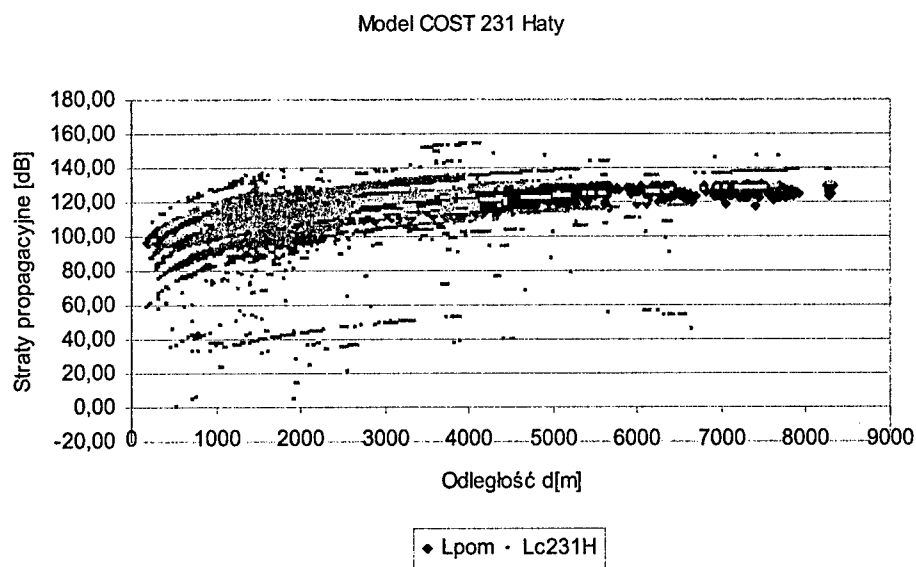
przy czym $a(h_a)$ wyraża się jako:

$$a(h_a) = (1,1 \lg(f[\text{MHz}]) - 0,7) h_a[m] - (1,56 \lg(f[\text{MHz}]) - 0,8). \quad (3.1)$$

3. ANALIZA PRZYDATNOŚCI

Podstawą opracowania nowej formy modelu COST 231 Haty były wyniki badań pomiarowych wykonanych w kilku dużych aglomeracjach miejskich, położonych w różnych częściach kraju. Łącznie wykonano 8556 pomiarów tłumienia propagacyjnego sygnału w radiowych łączach dostępowych, pochodzącego od dziewięciu różnych stacji bazowych, zlokalizowanych w: Gdańsku (GDA), Gdyni (GDY), Krakowie (KRA) i Poznaniu (POZ).

W pierwszej części badań wyniki tych pomiarów porównano z wartościami tłumienia obliczonymi przy użyciu klasycznej formy badanego modelu, wyrażonej zależnościami (1) i (1.1). Na tej podstawie obliczono błędy: średni ME (*mean error*) i średniokwadratowy MSE (*mean square error*), pomiędzy wartościami pomierzonymi i obliczonymi, łącznie dla wszystkich pomiarów, jak i również oddzielnie dla pomiarów od poszczególnych stacji bazowych. Rozrzut przestrzenny wartości tłumienia pomierzonych i obliczonych, w funkcji długości trasy propagacji, przedstawiono na rys. 2.



Rys. 2. Rozrzut przestrzenny wartości tłumienia pomierzonych i obliczonych przy użyciu klasycznej formy modelu COST 231 Haty – wyrażenia (1) i (1.1)

W rezultacie łączny błąd średni pomiędzy wartościami pomierzonymi i obliczonymi przy użyciu klasycznej formy modelu wyniósł 3,7 dB, z czym był związany błąd średniokwadratowy równy 15,7 dB. Błędy te w odniesieniu do poszczególnych stacji bazowych zebrano w tabeli 1.

Następnie, na podstawie tych błędów przeprowadzono wspomnianą analizę regresji liniowej funkcji błędu z dwiema zmiennymi niezależnymi, otrzymując w rezultacie nową postać badanego modelu, opisaną zależnościami (3) i (3.1).

W celu określenia przydatności nowej formy tego modelu, porównano wyniki pomiarowe z wynikami obliczeń tłumienia propagacyjnego wykonanymi przy użyciu formy modelu zmodyfikowanej dla warunków propagacyjnych w systemach dostępowych. Otrzymany w ten sposób łączny błąd średni wyniósł tylko 0,004 dB. Natomiast związany z nim błąd średniokwadratowy był równy 3,385 dB. Jest to zdecydowanie lepiej aniżeli to było dla klasycznej formy modelu. Błędy te dla poszczególnych stacji bazowych zebrano w tabeli 2.

Tabela 1

Zestawienie błędów: średnich ME oraz średniokwadratowych MSE dla poszczególnych stacji bazowych – klasyczna forma modelu COST 231 Haty, wyrażenia (1) i (1.1)

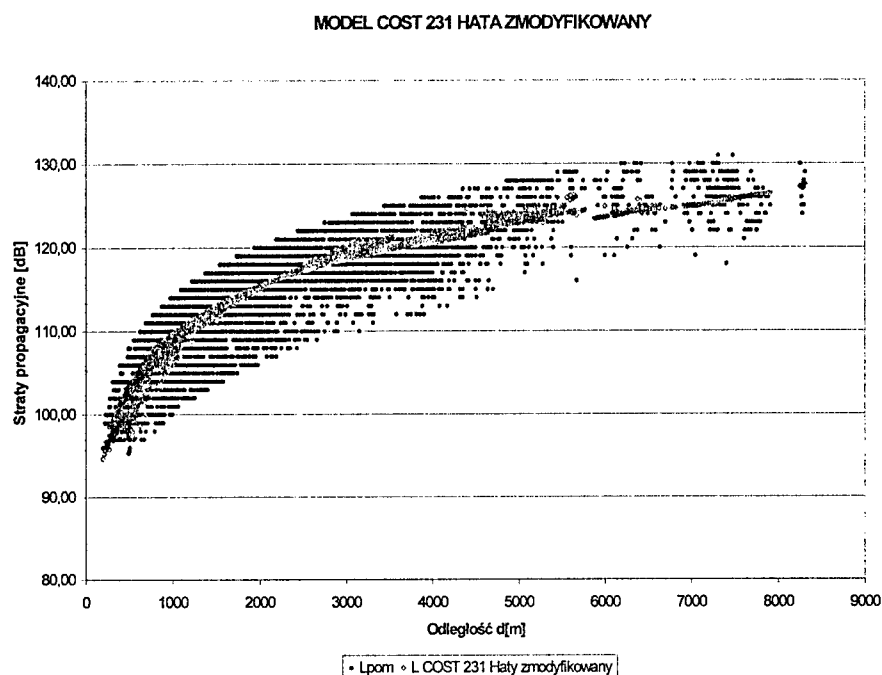
Miasto – numer stacji	ME [dB]	MSE [dB]	Liczebność pomiarów
GDA01	-0,555	13,931	2068
GDY01	10,881	22,551	265
GDY02	-0,154	15,472	466
GDY05	1,622	21,522	499
KRA01	2,218	20,590	1392
KRA02	5,080	8,902	961
KRA03	5,781	9,344	716
POZ01	8,139	15,159	1358
POZ02	7,949	15,420	831

Tabela 2

Zestawienie błędów: średnich ME oraz średniokwadratowych MSE dla poszczególnych stacji bazowych – zmodyfikowana forma modelu COST 231 Haty, wyrażenia (3) i (3.1)

Miasto – numer stacji	ME [dB]	MSE [dB]	Liczebność pomiarów
GDA01	-0,443	2,816	2068
GDY01	-1,383	3,931	265
GDY02	-0,508	4,182	466
GDY05	0,096	3,134	499
KRA01	1,571	3,134	1392
KRA02	0,986	3,116	961
KRA03	1,724	3,169	716
POZ01	-1,148	3,950	1358
POZ02	-1,281	4,237	831

Natomiast rozrzut przestrzenny wartości tłumienia: pomierzonych i obliczonych przy użyciu zmodyfikowanej formy modelu COST 231 Haty przedstawiono na rys. 3.



Rys. 2. Rozrzut przestrzenny wartości tłumienia pomierzonych i obliczonych przy użyciu zmodyfikowanej formy modelu COST 231 Hata – wyrażenia (3) i (3.1)

4. PODSUMOWANIE

Przeprowadzona analiza porównawcza pokazuje dużą przydatność nowej formy modelu COST 231 Hata do warunków występujących w radiowych sieciach dostępowych, pracujących w terenie zabudowanym.

Określony eksperymentalnie zakres stosowalności opracowanej formy tego modelu jest następujący:

- długości d tras propagacyjnych od 0,2 do 8,3 km,
- znormalizowana wysokość zawieszenia anten stacji bazowych od 1 do 106 m,
- wysokość zawieszenia anteny stacji abonenckiej 3 m.

Reasumując, na podstawie przeprowadzonych badań można stwierdzić, że zaproponowana postać modelu COST 231 Hata nadaje się do projektowania radiowych systemów dostępowych, przeznaczonych do pracy w warunkach miejskich.

BIBLIOGRAFIA

- [1] Katulski R.J., Kiedrowski A.: *Analiza projektowa radiowej stacji dostępowej w terenie zabudowanym*, Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej, seria: Technologie Informacyjne, 1/2003.
- [2] Katulski R.J., Kiedrowski A.: *Model COST 231 Hata w warunkach pracy radiowego systemu dostępowego w mieście*, (materiały) Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, KKRRiT – Wrocław, czerwiec 2003.

NEW FORM OF THE COST 231 HATA MODEL**Summary**

The problem of the radio-access propagation path loss in urban area by use the COST 231 Hata model has been analysed. The new form of the model is proposed. The usefulness of the form for the access systems design procedures has been concluded.

Ryszard J. Katulski, Michał Mikołajski

Katedra Systemów i Sieci Radiokomunikacyjnych, Politechnika Gdańska

UKŁADY ANTENOWE MIMO W TELEKOMUNIKACJI BEZPRZEWODOWEJ

Streszczenie

Przedmiotem pracy są podstawy techniki zbiorczego nadawania i odbioru dla przyszłych systemów telekomunikacji bezprzewodowej. Przedstawiono genezę powstania koncepcji MIMO (*Multiple Input Multiple Output*) oraz opisano w sposób systemowy tor transmisyjny rozwiązania użytkowego tej koncepcji. Scharakteryzowano działanie poszczególnych bloków tego toru, w szczególności podano macierze sygnałowe różnych rodzajów modulatora przestrzenno-czasowego. Dokonano analizy wyników przeprowadzonych badań symulacyjnych tych modulatorów. W zakończeniu oceniono ich przydatności w systemach telekomunikacji ruchomej.

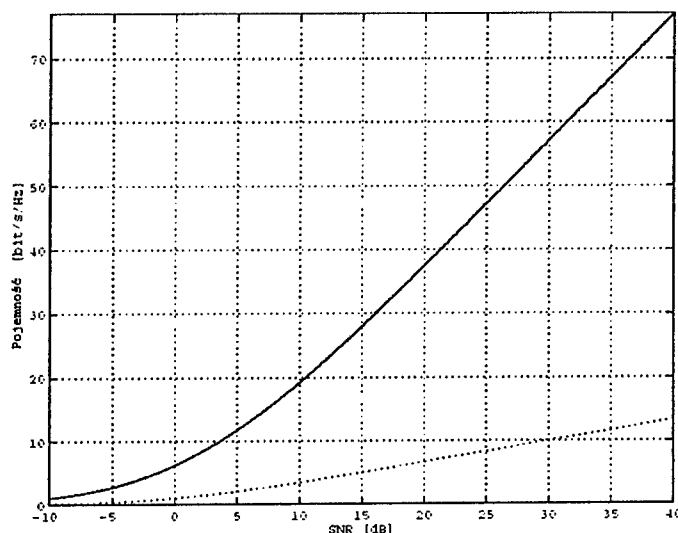
1. WSTĘP

Systemy MIMO (*Multiple Input Multiple Output*) należą do jednej z najbardziej rozwojowych dziedzin badawczych w obszarze komunikacji bezprzewodowej. Jest to spowodowane faktem oferowania przez kanał radiowy typu MIMO znaczącego wzrostu pojemności radiowego systemu bezprzewodowego, w stosunku do klasycznego kanału typu SISO (*Single Input Single Output*). Ten wzrost efektywności widmowej w kanale MIMO wynika z zastosowania zróżnicowanych przestrzennie różnych dróg propagacji sygnałów radiowych, widzianych zarówno od strony nadajnika jak i odbiornika. Podejście takie jest możliwe dzięki zastosowaniu systemów wieloantenowych. W takich warunkach strumień danych generowany przez pojedynczego użytkownika jest demultipleksowany na n_T podstrumieni nadawanych przez poszczególne anteny w układzie wieloantenowym, gdzie n_T jest równe liczbie anten nadawczych. Sygnały te mogą być odbierane w ogólności przez n_R anten odbiorczych.

Z uwagi na trudności w stosowaniu wielu anten odbiorczych w przenośnym terminalu użytkownika końcowego, w praktyce pod uwagę brane są rozwiązania typu MISO (*Multiple Input Single Output*), w których to ograniczono się do stosowania układu wieloantenowego jedynie w stacji bazowej transmitującej dane w kierunku użytkownika ruchomego.

2. EFEKTYWNOŚĆ WIDMOWA SYSTEMU MIMO

Zagadnienie poprawy efektywności widmowej osiąganej w systemach MIMO, w porównaniu do systemów SISO, zilustrowano na rys. 1. Metodę zastosowaną do jego uzyskania prześledzić można w [1]. W pracy [2] wykazano, że w określonych warunkach pojemność systemu wzrasta liniowo wraz z $\min(n_T, n_R)$. Ponieważ w warunkach rzeczywistych istnieje pewna korelacja pomiędzy sygnałami nadawanymi z poszczególnych anten, tak więc w praktyce zwiększanie pojemności nie będzie tak optymistyczne.



Rys.1. Pojemność Shannona kanału SISO (linia kropkowana) w porównaniu do ergodycznej pojemności kanału MIMO (linia ciągła), przy $n_T=n_R=6$

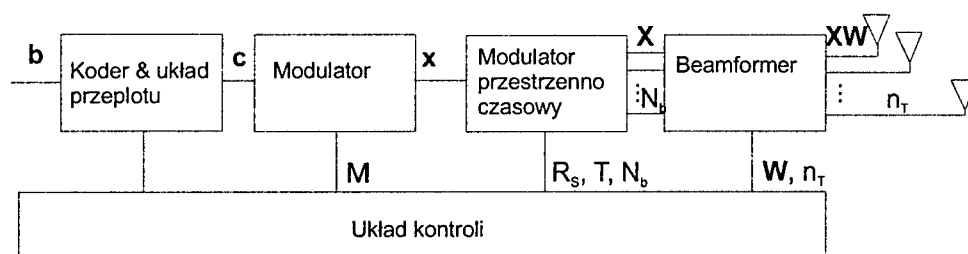
Podstawowe techniki zwielokrotnionego nadawania i odbioru znane były już w latach osiemdziesiątych [3]. Obecne zainteresowanie techniką MIMO datuje się od opublikowania raportu badawczego [2], w którym zasygnalizowano możliwość znacznie lepszego wykorzystania kanału radiowego niż miałyby to miejsce przy stosowaniu klasycznych metod zwielokrotnienia. W ramach prac nad systemami czwartej generacji założono, że dostępne dla użytkownika przepływności powinny się zawierać w przedziale od 100Mbps do 1Gbps, zajmując przy tym pasmo o szerokości co najmniej 100MHz. Jednym ze sposobów osiągnięcia w praktyce tego założenia jest zastosowanie minimum czterech anten nadawczych i tyle samo anten odbiorczych oraz modulacji 16QAM [4].

3. TOR TRANSMISYJNY W SYSTEMIE MIMO (MISO)

Omawianą w niniejszej pracy koncepcję techniki MIMO można realizować w dwojaki sposób, tzn. w systemie z pętlą otwartą lub zamkniętą. Różnica polega na tym, że przy pracy z pętlą zamkniętą powstaje sprzężenie zwrotne pomiędzy stacją bazową i stacją

ruchomą, służące do przesyłania informacji o stanie kanału radiowego. Pozwala to odpowiednio modyfikować nadawany sygnał, tak aby osiągnąć dopasowanie aktualnej charakterystyki modulatora przestrzenno-czasowego oraz układu kształtowania właściwości kierunkowych (beamformera) do warunków propagacji sygnału w kanale radiowym. Przedmiotem niniejszej pracy są różne odmiany modulatorów przestrzenno-czasowych pracujących w pętli otwartej.

W ogólności układ blokowy części nadawczej toru transmisyjnego w systemie MIMO, co odnosi się także do rozwiązania typu MISO, przedstawiono na rys. 2:



Rys.2. Tor transmisyjny w systemie MIMO (MISO)

Tor ten składa się z następujących bloków, które realizują określone funkcje systemowe:

- Koder wraz układem przeplotu, do którego wpływa strumień bitów ze źródła informacji. Na wyjściu tego układu otrzymujemy wektor bądź strumień zakodowanych bitów c .
- Strumień c jest następnie podawany do układu modulatora, który mapuje M bitów tworząc zespolony wektor zmodulowany x , w postaci strumienia symboli.
- Z kolei strumień symboli trafia do modulatora przestrzenno-czasowego, który sekwencje $R_s T$ symboli mapuje w macierz X o rozmiarze $T \times N_b$, przygotowując w ten sposób symbole do transmisji przez T ortogonalnych zasobów transmisyjnych (np. czasu, kodów rozpraszających, podnośnych) w ramach N_b wiązek promieniowania układu antenowego.
- Wyjście modulatora przestrzenno-czasowego zasila obwód beamformera, który konstruuje N_b wiązek promieniowania i wysterowuje nimi n_T zasobów nadawczych (anten). Działania wykonywane przez beamformer opisane są przez macierz W .

4. OPIS MACIERZOWY MODULATORA PRZESTRZENNO-CZASOWEGO

Zapis macierzowy odnosi się zmapowanej formy sygnałów na wyjściu modulatora przestrzenno-czasowego. Poniżej przedstawiono podstawowe formy takiego zapisu, które wynikają z koncepcji działania systemu MIMO. W zapisach tych przyjęto, że: częstotliwość symboli $R_s = 1$ oraz $n_T = 4$ anteny, przy czym $T = 4$ okresom symboli. Zapisy te ilustrują sposoby, przy użyciu których osiąga się zamierzone dla układów MIMO efekty trans-

misyjne. Są to następujące koncepcje modulatora przestrzenno-czasowego, określone skrótowymi nazwami:

- modulator typu STTD-OTD (*Space Time Transmit Diversity with Orthogonal Transmit Diversity*),
- modulator typu PSK-rank,
- modulator cykliczny,
- modulator warstwowy,
- modulator typu ABBA.

STTD-OTD

Jest to liniowy modulator ortogonalny z suboptymalnym różnicowaniem. Najprostsza macierz modulatora STTD dla dwóch anten nadawczych i jednej anteny odbiorczej przyjmuje postać podaną przez Alamoutiego w [5]:

$$\mathbf{X}_{(x_1, x_2)} = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}. \quad (1)$$

Dokonując połączenia dwóch kopii (1), otrzymujemy blokowo-diagonalną macierz:

$$\mathbf{X}_{STTD-OTD} = \begin{bmatrix} \mathbf{X}_A(x_1, x_2) & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{X}_B(x_3, x_4) \end{bmatrix}. \quad (2)$$

w której: $\mathbf{X}_A(x_1, x_2)$, $\mathbf{X}_B(x_3, x_4)$ – macierze STTD Alamoutiego dla sygnałów (x_1, x_2) oraz sygnałów (x_3, x_4) , przy czym $\mathbf{0}_2$ stanowi macierz zer o wymiarach 2×2 .

PSK-rank

Macierz modulatora PSK-rank jest przykładem macierzy pełnego rzędu. W schemacie 4×4 przyjmuje ona poniższą postać:

$$\mathbf{X}_{PSK-rank} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_1 + x_2 \\ x_3 & x_4 & x_1 + x_2 & x_2 + x_3 \\ x_4 & x_1 + x_2 & x_2 + x_3 & x_3 + x_4 \end{bmatrix}. \quad (3)$$

Dokładną analizę powstania takiej macierzy zawiera praca [6].

Modulator cykliczny

Natomiast macierz modulatora cyklicznego przyjmuje następującą postać:

$$\mathbf{X}_{\text{cykl}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_1 & x_2 & x_3 \\ x_3 & x_4 & x_1 & x_2 \\ x_2 & x_3 & x_4 & x_1 \end{bmatrix}. \quad (4)$$

Modulator warstwowy - schemat quasi-ortogonalny

Modulator taki został przedstawiony szczegółowo w pracy [7]. Jego macierz określona jest następująco:

$$\mathbf{X}_{3+1} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2^* & x_1^* & x_4 & -x_3 \\ -x_3^* & -x_4 & x_1^* & x_2 \\ -x_4 & x_3^* & -x_2^* & x_1 \end{bmatrix} \quad (5)$$

ABBA

Powyższy skrót pochodzi od rozmieszczenia w macierzy modulatora submacierzy A oraz B. Z matematycznego punktu widzenia jest to schemat 2+2 warstwowy quasi-ortogonalny. Szczegółowy opis można znaleźć w pracy [8]. Macierz tego modulatora ma postać:

$$\mathbf{X}_{ABBA}(x_1, x_2, x_3, x_4) = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2^* & x_1^* & -x_4^* & x_3^* \\ x_3 & x_4 & x_1 & x_2 \\ -x_4^* & x_3^* & -x_2^* & x_1^* \end{bmatrix} \equiv \begin{bmatrix} \mathbf{X}_A & \mathbf{X}_B \\ \mathbf{X}_B & \mathbf{X}_A \end{bmatrix} \quad (6)$$

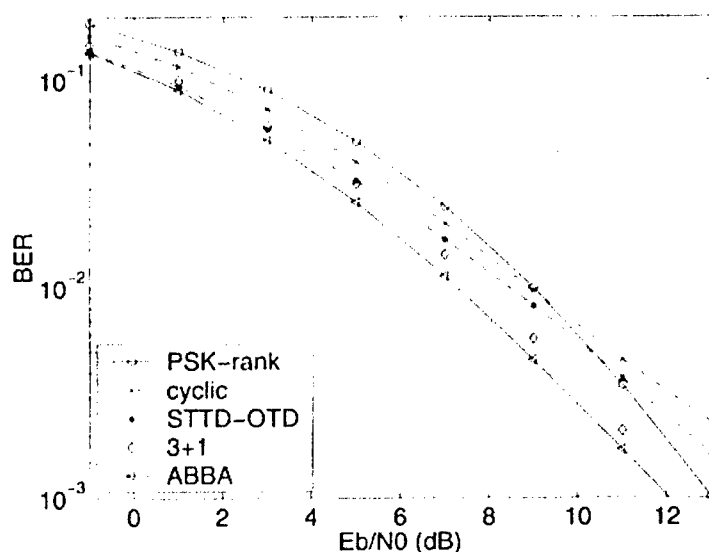
5. BADANIA SYMULACYJNE

Do badań symulacyjnych opracowano model kanału oraz modele modulatorów MIMO w postaci modułów Matlaba. Symulacji dokonano przy następujących założeniach:

- cztery anteny nadawcze
- zaniki bloków sygnału zgodne z rozkładem Rayleigha
- długość bloku równa cztery

- jedna antena odbiorcza
- idealna estymacja parametrów kanału w odbiorniku
- detekcja sygnałów za pomocą algorytmu ML (*Maximum Likelihood*)

W wyniku przeprowadzonych badań symulacyjnych opracowano wykres przedstawiony na rys. 3, który obrazuje zależność stopy błędów BER od stosunku E_b/N_0 .



Rys. 3. Stopa błędów BER w funkcji E_b/N_0 dla wybranych modulatorów czasowo-przestrzennych, przy zastosowaniu czterech anten nadawczych

Porównując otrzymane charakterystyki można sformułować następujące wnioski:

- Różnice pomiędzy poszczególnymi modulatorami wynikają głównie z interferencji własnych, które modulator wnosi do kanału radiowego. Im większa jest ich liczba, tym mniejsze są możliwości różnicowania drogi sygnału;
- Dobrze skonstruowany nieortogonalny schemat modulacji przestrzenno-czasowej typu ABBA posiada lepsze właściwości niż ortogonalny schemat STTD-OTD, we wszystkich zakresach stosunku E_b/N_0 . Jest to spowodowane tym, że ABBA zapewnia pełną ochronę przed błędami o małym dystansie euklidesowym;
- Właściwości modulatora przestrzenno-czasowego dla średnich wartości SNR mogą zostać poprawione poprzez zastosowanie macierzy, które w maksymalny sposób różnicują kolejne symbole. Oznacza to, że transmisja każdego symbolu jest w równym stopniu podzielona pomiędzy wszystkie anteny, z jednoczesnym użyciem różnych okresów symbolu dla różnych anten.

6. PODSUMOWANIE

W wyniku prowadzonych prac badawczych nad systemami trzeciej i czwartej generacji powstała całkowicie nowa koncepcja zwiększenia możliwości transportowych kanału bezprzewodowego, poprzez wykorzystanie jego właściwości statystycznych. W wyniku tego oczekiwane jest znaczne zwiększenie wydajności widmowej i osiągnięcie przez systemy bezprzewodowe przepływności oferowanej dotychczas jedynie przez techniki światłowodowe oraz sieci Ethernet. Pełne wykorzystanie możliwości płynących z techniki MIMO ograniczone jest liczbą anten odbiorczych zastosowanych w terminalu ruchomym. Oczekuje się, że w przyszłych systemach radiowych wykorzystane zostaną częstotliwości z zakresu pasma znacznie wyższego niż ma to miejsce w chwili obecnej. Pozwoli to na efektywną miniaturyzację anten oraz pełne wykorzystanie zalet techniki MIMO, bez potrzeby rezygnacji z miniaturyzacji terminali ruchomych.

Odbiór i nadawanie zbiorcze posiadają znaczny potencjał rozwojowy. Właściwości statystyczne kanału zależą jednak w znacznym stopniu od środowiska propagacji sygnału radiowego. Dlatego też nie wszystkie środowiska propagacyjne są szczególnie predestynowane do stosowania opisanych w pracy rozwiązań modulatora przestrzenno-czasowego. Najlepszym, rokującym największe nadzieje środowiskiem są mikrokomórki we wnętrzach budynków. Tuż za nimi lokują się komórki miejskie, szczególnie w miastach o wysokim stopniu rozwoju i rozbudowanej architekturze. Najgorszym środowiskiem propagacyjnym jest środowisko wiejskie, a to z uwagi na duży stopień skorelowania pomiędzy sygnałami w poszczególnych ścieżkach propagacji. Ograniczenia te silnie wpływają na tendencje rozwojowe systemów MIMO. Oczekuje się, że pierwszym obszarem zastosowań będą sieci WLAN o zasięgu stacji bazowej do 50 metrów, oferujące efektywną przepływność 1 Gbps.

BIBLIOGRAFIA

- [1] Holter B.: *On the capacity of the MIMO channel – A tutorial introduction*, http://www.tele.ntnu.no/projects/beats/Documents/MIMO_introduction.pdf, 2001.
- [2] Telatar I.: *Capacity of multiantenna gaussian channels* AT&T technical memorandum, 1995.
- [3] Winters J.: *On the capacity of radio communication systems with diversity in a Rayleigh fading environment*, IEEE J. Sel. Areas Commun., Vol. 5, no. 5, pp. 871-878, June 1987.
- [4] Hottinen A., Tirkkonen O., Wichman R.: *Multi-antenna Transceiver Techniques for 3G and Beyond*, Wiley 2003.
- [5] Alamouti S.: *A simple transmitter diversity scheme for wireless communications*, IEEE J. Sel. Areas Commun., Vol. 16, pp. 1451-1458, Oct. 1998.
- [6] Hammons A., El Gamal H.: *On the theory of space time codes for PSK modulation*, IEEE Trans. Inf. Th., Vol. 46, no. 2, pp. 524-542, 2000.
- [7] Tirkkonen O.: *Layered space-time block codes*, manuscript, 2001.
- [8] Tirkkonen O., Boariu A., Hottinen A.: *Minimal non-orthogonality rate one space time block code for 3+ Tx antennas*, IEEE International Symposium on Spread spectrum Techniques and Applications, Sept. 2000, Vol. 2, pp. 429-432.

MIMO ANTENNA ARRAYS IN WIRELESS COMMUNICATION**Summary**

This article describes the topic of Transmit and Receive Diversity in the scope of future mobile communication systems. Paper begins with the genesis of MIMO (Multiple Input Multiple Output) concept and the characteristics of representative solutions from the system point of view. A methodology of system analysis is presented and some simulation results are given. Finally, a brief conclusion about MIMO feasibility is discussed.

Ryszard J. Katulski, Jarosław Sadowski

Katedra Systemów i Sieci Radiokomunikacyjnych, Politechnika Gdańska

INTERAKCJA ELEKTRODYNAMICZNA ANTENY Z UŻYTKOWNIKIEM TELEFONU KOMÓRKOWEGO

Streszczenie

W pracy przedstawiono zagadnienie modelowania rozkładu pola elektrycznego w polu bliskim anteny telefonu komórkowego, z punktu widzenia badania zjawiska pochłaniania energii tego pola przez elementy ciała użytkownika telefonu. Opisano taki model i omówiono wyniki przeprowadzonych badań symulacyjnych, w zakresie odnoszącym się do pochłaniania energii przez dłoń operatora trzymającego aktywny telefon komórkowy oraz określono wpływ tej dłoni na odkształcenie właściwości kierunkowych anteny.

1. WPROWADZENIE

Coraz powszechniejsze stosowanie techniki radiowej w systemach telekomunikacyjnych powiększa stopień oddziaływania pola elektromagnetycznego na użytkowników tych systemów, co szczególnie odnosi się do operatorów telefonów komórkowych.

Problem wzajemnego oddziaływania – interakcji elektrodynamicznej – urządzenia antenowego i ciała użytkownika telefonu komórkowego można rozpatrywać w sposób dwojaki, tzn. jako:

- wpływ wytwarzanego pola elektromagnetycznego na organizm ludzki, w wyniku prądów indukowanych w tkankach i płynach ustrojowych człowieka, oraz
- oddziaływanie na właściwości elektryczne anteny wtórnego pola elektromagnetycznego wytwarzanego przez te prądy.

W pierwszym ujęciu, na podstawie wielkości energii pochłanianej przez napromieniowywane fragmenty ciała ludzkiego, określa się lokalne przyrosty temperatury, na podstawie czego można wnioskować co do skutków fizjologicznych takiego promieniowania. W drugim ujęciu, poprzez określenie wielkości elektrodynamicznego sprzężenia zwrotnego, można określić zmianę właściwości kierunkowych oraz impedancyjnych anteny, co ma bezpośredni wpływ na funkcjonowanie łącza radiokomunikacyjnego.

Jest oczywiste, że oba wymienione aspekty interakcji elektromagnetycznej są ze sobą ściśle powiązane, co jest przedmiotem intensywnych badań, szczególnie w kontekście używania telefonów komórkowych. Badania takie często dotyczą typowych konfiguracji układu: użytkownik – telefon komórkowy, np. aparat trzymany w ręku przy głowie

człowieka, w trakcie rozmowy. Zazwyczaj w badaniach tych analizuje się wpływ pola elektromagnetycznego na głowę operatora. Natomiast nie bierze się pod uwagę sytuacji nietypowych, charakterystycznych dla stopnia dużego zagrożenia, np. chwytanie aparatu w części antenowej jako reakcja na odbiór przychodzącego wywołania, kiedy to zagrożona jest dłoń człowieka.

Przedmiotem artykułu jest analiza wpływu promieniowania elektromagnetycznego anteny telefonu komórkowego na dłoń użytkownika tego telefonu oraz wpływ obecności tego użytkownika na właściwości kierunkowe anteny. Badania te zostały przeprowadzone przy użyciu symulacji komputerowych, w paśmie 1800 MHz.

2. OPIS MODELU SYMULACYJNEGO

2.1. Podstawy analityczne

Podstawową wielkością stosowaną do określenia narażenia organizmu operatora telefonu komórkowego na działanie pola elektromagnetycznego jest tzw. SAR (*Specific Absorption Rate*), którego wartość wyznacza się na podstawie rozkładu pola elektrycznego E zaindukowanego w tkance człowieka, co można zapisać w poniższy sposób:

$$SAR = \frac{\sigma \cdot E^2}{\rho}, \quad (1)$$

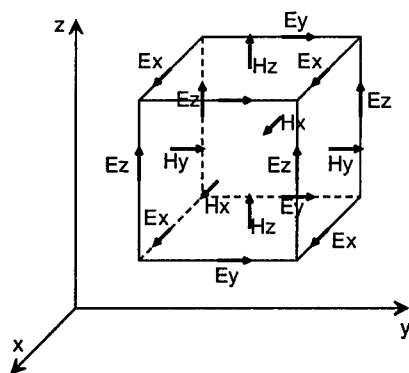
w którym: σ oraz ρ oznaczają odpowiednio, konduktywność oraz gęstość napromieniowanej tkanki.

Z kolei do określenia rozkładu tego pola w ośrodku silnie niejednorodnym, jakim w ogólności jest tkanka człowieka, powszechnie stosuje się metodę *FDTD* (*Finite Difference Time Domain*), np. [1]. Metoda ta polega na zastąpieniu występujących w równaniach Maxwella pochodnych cząstkowych przez przyrosty skończone, zarówno w dziedzinie czasu jak i w przestrzeni.

Opracowując model numeryczny analizowanego problemu przyjęto poniższe oznaczenie na wartość funkcji $F^n(i, j, k)$, opisującej dowolną składową pola elektromagnetycznego ($E_x, E_y, E_z, H_x, H_y, H_z$) w trójwymiarowym elemencie tkanki – patrz rys. 1, określonym przez stany liczników (i, j, k) , oraz w przedziale czasu określonym przez stan licznika n , tzn.:

$$F^n(i, j, k) = F(i \cdot \delta x, j \cdot \delta y, k \cdot \delta z, n \cdot \delta t), \quad (2)$$

gdzie stany liczników (i, j, k) oraz n są określone przez liczby całkowite $(0, 1, 2, \dots)$, wyznaczające kolejne kroki liczenia, przy czym $\delta x, \delta y, \delta z$ oznaczają przyrosty długości – kroki kwantyzacji przestrzennej względem odpowiednich osi, zaś δt oznacza przyrost czasu.



Rys. 1. Element tkanki w trójwymiarowym układzie współrzędnych

W wyniku rozwiązania równań Maxwella otrzymuje się układ sześciu równań algebraicznych, które w przystępny sposób można zaimplementować komputerowo. Przykładowo, równanie opisujące składową E_x wyraża się następująco:

$$\begin{aligned}
 E_x^{n+1}\left(i+\frac{1}{2}, j, k\right) &= \left(1 - \frac{\sigma\left(i+\frac{1}{2}, j, k\right) \cdot \delta t}{\varepsilon\left(i+\frac{1}{2}, j, k\right)}\right) \cdot E_x^n\left(i+\frac{1}{2}, j, k\right) + \\
 &+ \frac{\delta t}{\varepsilon\left(i+\frac{1}{2}, j, k\right) \cdot \delta x} \cdot \left[H_z^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j+\frac{1}{2}, k\right) - H_z^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j-\frac{1}{2}, k\right) + \right. \\
 &\left. + H_y^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j, k-\frac{1}{2}\right) - H_y^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j, k+\frac{1}{2}\right) \right].
 \end{aligned} \quad (3)$$

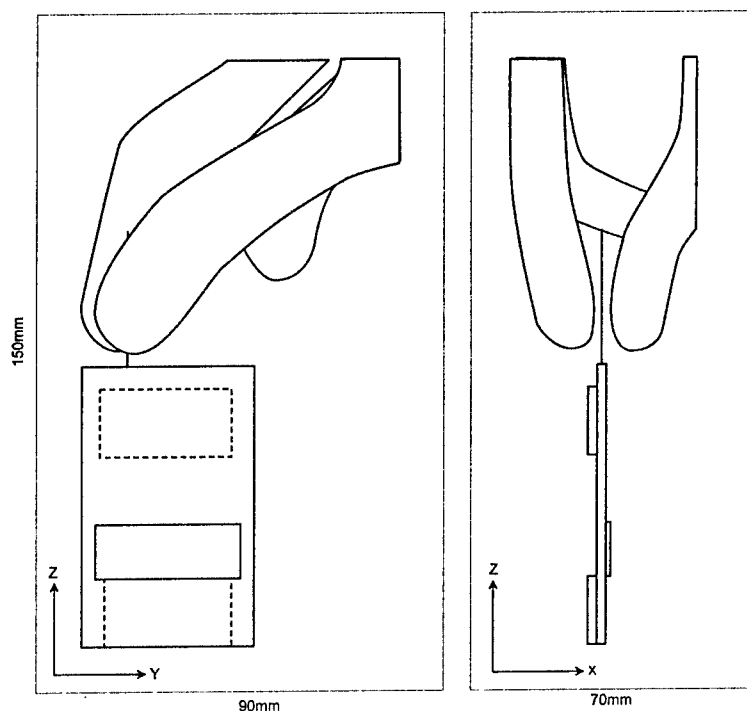
Ze względu na ograniczone zasoby pamięci i mocy obliczeniowej komputerów, symulowaną przestrzeń biologiczną wybranej części ciała ludzkiego, opisaną układem równań Maxwella, należy otoczyć warstwą nie powodującą odbicia od granicy tej przestrzeni. Powszechnie do tego celu stosuje się warstwę o skrótowej nazwie *PML* (*Perfectly Matched Layer*), z odpowiednio dobranym profilem konduktywności, zapewniającym silne tłumienie padających na nią fal elektromagnetycznych.

Opisana powyżej metoda pozwala określić rozkład składowych pola elektromagnetycznego wewnątrz badanej części ciała ludzkiego, co z kolei umożliwia obliczenie wynikającego z tego rozkładu SAR. Ponadto, znając rozkład pola w strefie bliskiej anteny

telefonu komórkowego, można obliczyć rozkład pola promieniowanego w strefie dalekiej i na tej podstawie wyznaczyć charakterystykę promieniowania tej anteny.

2.2. Geometria badanego obszaru

Do badań symulacyjnych wybrano obszar o wymiarach $70 \times 90 \times 150$ mm, otoczony zamkniętą warstwą PML o grubości 8δ – tzn. odpowiednio $8\delta_x$, $8\delta_y$ oraz $8\delta_z$, z wykładniczo zmienną konduktywnością. Geometrię tego obszaru dla przypadku telefonu komórkowego trzymanego za antenę palcami prawej dłoni, pokazano na rys. 2.



Rys. 2. Geometria badanego obszaru

Model elektryczny telefonu komórkowego składał się z zestawu czterech prostopadłościów przewodzących, odzwierciedlających jego konstrukcję, a więc płytę czołową, osłony ekranujące bloki składowe i akumulator, wraz z anteną prętową pracującą w paśmie 1800 MHz. Krok kwantyzacji przestrzennej przyjęto równy 1 mm, względem każdej z osi, tzn.: $\delta x = \delta y = \delta z = 1$ mm.

Badany materiał biologiczny scharakteryzowano następującymi wielkościami:

- obszar kostny poprzez $\epsilon_r = 11,8$ oraz $\sigma = 0,33$ S/m,
- obszar tkanki mięśniowej poprzez $\epsilon_r = 56,7$ oraz $\sigma = 1,92$ S/m.

Ze względu na duże wymagania metody FTDT, zarówno co do obszaru pamięci operacyjnej jak i koniecznego czasu obliczeń, badania symulacyjne wykonano przy użyciu

równoległej architektury komputera wieloprocesowego. Program obliczeniowy został napisany w języku C^{++} , z wykorzystaniem biblioteki *MPI*, co w razie konieczności pozwoli na implementację przy użyciu innego rodzaju komputerów.

3. WYNIKI BADAŃ SYMULACYJNYCH

Opisane poniżej badania symulacyjne wykonano w Centrum Informatycznym *Trójmiejskiej Akademickiej Sieci Komputerowej (TASK)*.

Badania te przeprowadzono w zakresie odnoszącym się do zjawiska pochłaniania energii pola elektromagnetycznego przez elementy dłoni użytkownika telefonu komórkowego oraz określenia wpływu obecności tej dłoni na właściwości kierunkowe anteny tego telefonu. Warunki energetyczne badań wynikały z mocy nadajnika telefonu komórkowego równej $0,25\text{ W}$, przy częstotliwości 1850 MHz . Telefon ten współpracował z antenami prętowymi o różnych długościach, tzn.: 26 lub 51 mm .

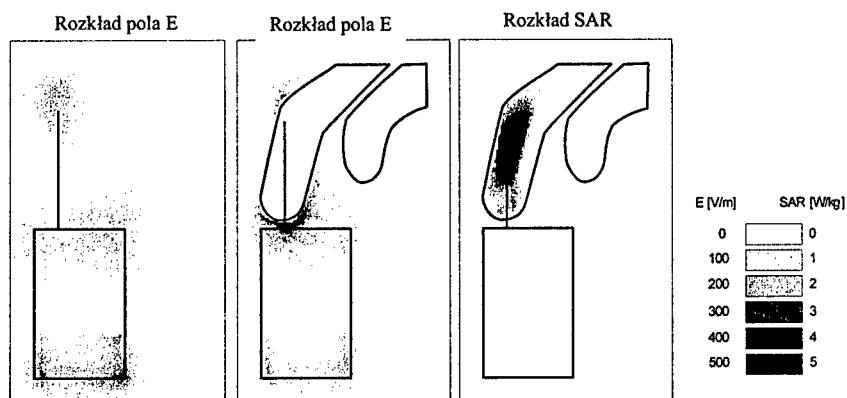
W badaniach tych symulowano dwa warianty modelu dłoni użytkownika:

- odwzorowanie uproszczone, ograniczone do jednorodnego obszaru o właściwościach elektrycznych jak dla tkanki mięśniowej, oraz
- odwzorowanie pełne, uwzględniające występowanie kości palców operatora.

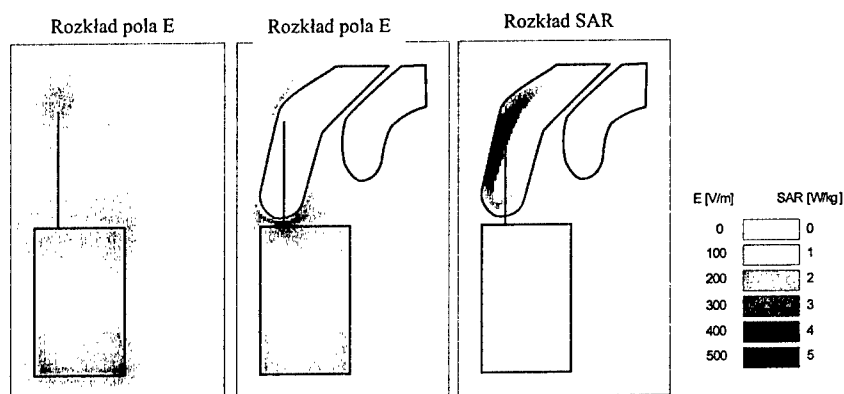
3.1. Pochłanianie energii pola elektromagnetycznego

Otrzymane rozkłady natężenia pola elektrycznego i *SAR*, wybrane przykładowo, przedstawiono na rys. 3-6, dla różnych długości anten telefonu komórkowego, obrazując oddziaływanie w przekroju kciuka lub palca wskazującego dłoni operatora – na rys. 2 w płaszczyźnie odpowiednio $x=20\text{ mm}$ lub $x=40\text{ mm}$. Każdy zobrazowany przypadek zawiera trzy sytuacje:

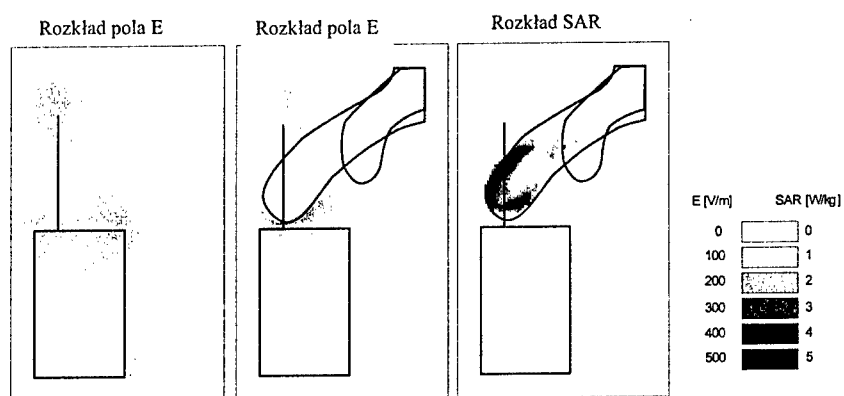
- rozkład pola elektrycznego w wybranym przekroju, pochodzącego od odosobnionego telefonu komórkowego z anteną prętową o określonej długości,
- rozkład pola elektrycznego od układu telefonu z dłonią użytkownika, oraz
- rozkład *SAR* od układu telefonu z dłonią użytkownika.



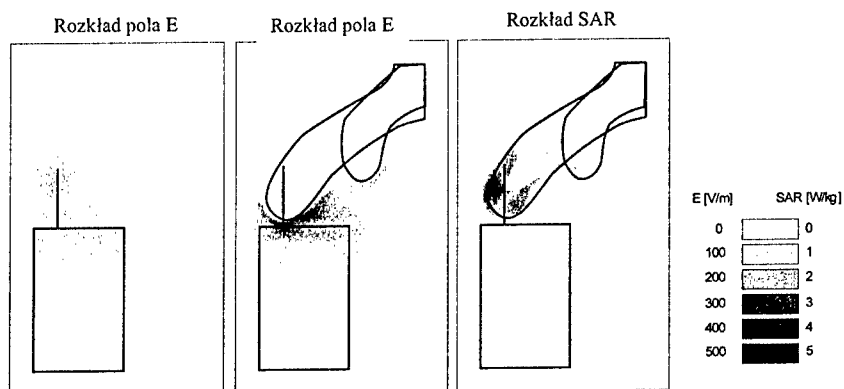
Rys. 3. Rozkłady pola elektrycznego i *SAR* w płaszczyźnie $x=20\text{ mm}$ (palec wskazujący)
– telefon komórkowy z dłuższą anteną, uproszczony model dłoni



Rys. 4. Rozkłady pola elektrycznego i SAR w płaszczyźnie $x=20\text{ mm}$ (palec wskazujący)
– telefon komórkowy z dłuższą anteną, pełny model dłoni



Rys. 5. Rozkłady pola elektrycznego i SAR w płaszczyźnie $x=40\text{ mm}$ (kciuk)
– telefon komórkowy z dłuższą anteną, pełny model dłoni



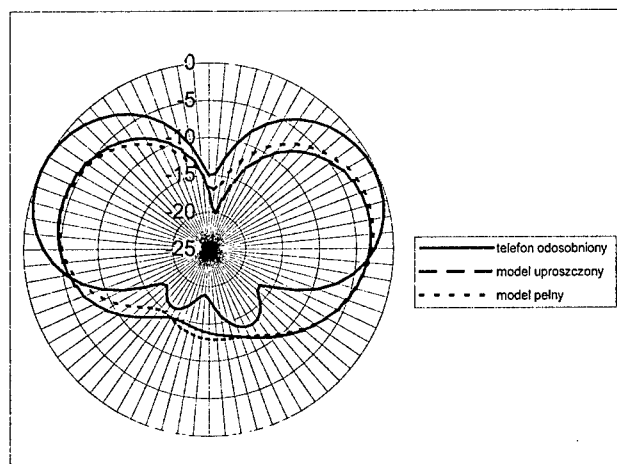
Rys. 6. Rozkłady pola elektrycznego i SAR w płaszczyźnie $x=40\text{ mm}$ (kciuk)
– telefon komórkowy z krótszą anteną, pełny model dłoni

Jak widać, w przypadku dłuższej anteny prętowej maksymalna wartość SAR w palcach dłoni użytkownika wyniosła 0,5 lub 15 W/kg, odpowiednio dla tkanki kostnej lub mięśniowej. Dla porównania, w przypadku zastosowania anteny krótszej, wartości te obniżyły się odpowiednio do 0,4 i 5 W/kg. Zalecenia międzynarodowe w zakresie ochrony przed promieniowaniem niejonizującym określają dopuszczalną wartość SAR na poziomie 2 W/kg, rozumianą jako średnią na każde 10 g masy ciała. Z przytoczonych powyżej wyników symulacji, uwzględniając ich przeliczenie z sześciątów o bokach 1 mm – przyjęty krok kwantyzacji, wynika przekroczenie dopuszczalnej wartości SAR w palcach dłoni użytkownika telefonu komórkowego.

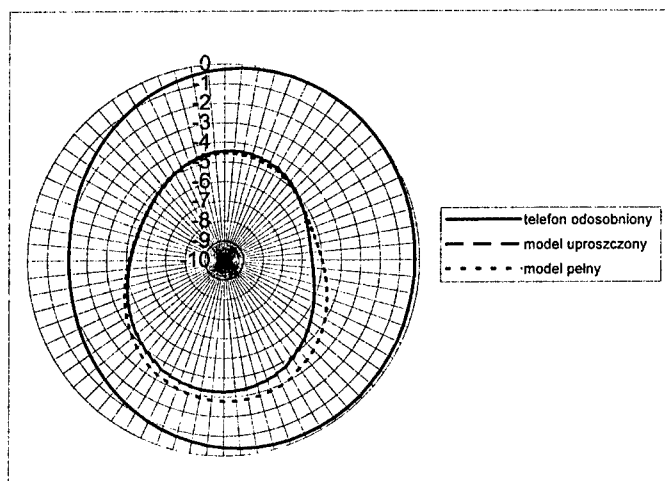
3.2. Odształcenie właściwości kierunkowych anteny

Charakterystyki promieniowania anteny telefonu komórkowego, wyliczone na podstawie komputerowych symulacji rozkładów natężenia pola elektrycznego, obrazują wpływ dłoni operatora na właściwości kierunkowe tej anteny.

Wybrane przykładowo przekroje tych charakterystyk w określonych płaszczyznach – patrz rys. 2 - przedstawiono na rys. 7-8, na każdym z nich zestawiono trzy charakterystyki promieniowania, odnoszące się do odosobnionego telefonu oraz dla zestawu: telefon – dłoń użytkownika, przy czym zastosowano uproszczony i pełny model tej dłoni.



Rys. 7. Charakterystyki promieniowania anteny telefonu komórkowego w płaszczyźnie X-Z, dla anteny dłuższej



Rys. 8. Charakterystyki promieniowania anteny telefonu komórkowego w płaszczyźnie X-Y, dla anteny dłuższej

Na uwagę zasługuje zobrazowanie porównawcze przedstawione na rys. 8, gdzie widać znaczące różnice pomiędzy poziomem energetycznym pola elektromagnetycznego promieniowanego przez antenę odosobnionego telefonu komórkowego i w sytuacji uwzględnienia dłoni operatora, co odbywa się kosztem energii pochłanianej przez jego dłoń.

PODSUMOWANIE

Przeprowadzone badania wyraźnie pokazują niekorzystny wpływ aktywnego telefonu komórkowego na dłoń użytkownika tego telefonu. Należy jednak podkreślić, że ma to charakter lokalny i może być znacznie ograniczone przez zastosowanie specjalnej konstrukcji urządzenia antenowego telefonu komórkowego.

BIBLIOGRAFIA

- [1] Wang J., Fujiwara O.: *Modelling of EM Interaction between a 5 GHz band Antenna mounted PC and a Human Body*, (materiały) Sixteenth Int. Wrocław Symposium on Electromagnetic Compatibility, EMC – Wrocław, czerwiec 2002.

EM INTERACTION BETWEEN ANTENNA AND PERSONAL TERMINAL USER

Summary

The problem of personal terminal antenna influences on a body of the terminal user have been analysed. The distributions of the SAR (Specific Absorption Rate) in a hand of the user, obtained by apply numerical simulation, are presented. The state of emergency is concluded.

Andrzej Kosiński

Zakład Łączności, Akademia Marynarki Wojennej

**PROMIENIE LASERA NIEBIESKO-ZIELONEGO
W ZASTOSOWANIU DO ŁĄCZNOŚCI Z OBIEKTAMI
ZANURZONYMI**

Streszczenie

W artykule zaprezentowano nowe kierunki badań i osiągnięcia w zakresie przekazywania informacji do zanurzonych okrętów podwodnych z zastosowaniem promieni laserowych. Omówiono właściwości propagacji promieni lasera w atmosferze i w wodzie morskiej. Scharakteryzowano wpływ ośrodka propagacji na wymagane parametry lasera. Przedstawiono problemy dotyczące transmisji informacji do okrętu podwodnego z wykorzystaniem promieni z zakresu widma niebiesko-zielonego.

1. WSTĘP

Rozwój konstrukcji laserów niebiesko-zielonych doprowadziła do jeszcze jednego ważnego ich zastosowania – do łączności z zanurzonymi okrętami podwodnymi. Efektywność działań bojowych atomowych okrętów podwodnych w znacznym stopniu zależy od szybkości i pewności funkcjonowania systemu łączności. Idealnym byłby taki system, który miałby możliwość przekazywania sygnałów przez wodę morską do głęboko zanurzonego okrętu, a ponadto aby transmisja odbywała się z dużymi szybkościami.

Już w końcu lat 70-tych w USA wraz z rozwojem systemu łączności konwencjonalnej tj. z wykorzystaniem fal elektromagnetycznych prowadzono prace badawcze pod nazwą „Laserowe systemy łączności z okrętami podwodnymi”. Wykorzystanie łącza optycznego jako alternatywnego w stosunku [1] do systemu łączności na falach elektromagnetycznych VLF i ELF spowodowane było obiecującymi cechami kanałów optycznych. Do podstawowych zalet telekomunikacyjnych kanałów optycznych należą – szybkość transmisji i skrytość łączności. Badano systemy łączności VLF i ELF oraz system łączności laserowej. Nieprzerwane przekazywanie sygnałów do zanurzonych okrętów podwodnych z wykorzystaniem bardzo długich fal elektromagnetycznych praktycznie jest możliwe niezależnie od czynników klimatycznych i meteorologicznych. Mankamentem tego sposobu przekazywania informacji jest mała szybkość transmisji – tym mniejsza im dłuższa fala. Zastosowanie [2] kanałów optycznych znacznie poprawia możliwości operacyjne systemu łączności.

2. PROPAGACJA PROMIENIA LASERA

Badając system łączności laserowej z zanurzonymi obiektami należy rozpatrzyć propagację wiązki w dwóch ośrodkach – w atmosferze i w wodzie morskiej. Tłumienie promieniowania w atmosferze jest zawsze rezultatem równocześnie występujących zjawisk absorpcji i rozproszenia. Największy wpływ wywiera środowisko na amplitudę. Osłabienie promieniowania elektromagnetycznego podlega prawu Bouguera-Lamberta-Beera :

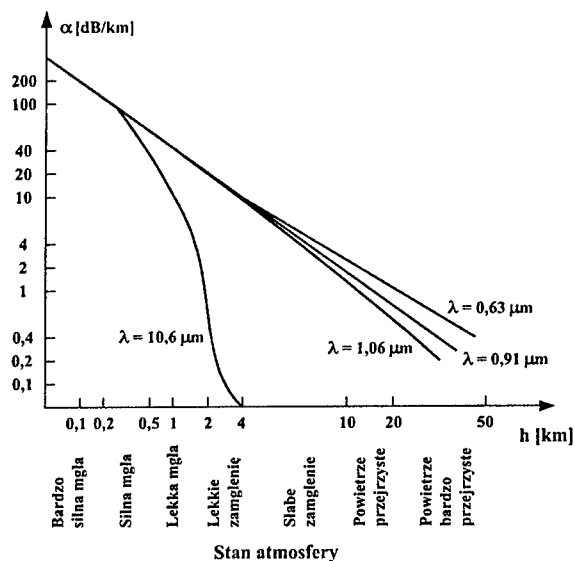
$$I = I_0 \exp(-\alpha R) \quad (2.1)$$

co oznacza, że jeżeli wiązka światła o natężeniu I_0 przechodzi przez ośrodek o grubości R to jej natężenie I zmaleje o wielkość $e^{-\alpha R}$, gdzie α oznacza współczynnik ekstynkcji, natomiast wielkość tę można nazwać transmisją środowiska, którą wyrażamy zależnością:

$$T_{\text{tr}} = I/I_0 = e^{-\alpha R} \quad (2.2)$$

Przy czym związek współczynnika ekstynkcji ze współczynnikiem absorpcji (α_a) i współczynnikiem rozproszenia (α_r) jest następujący:

$$\alpha = \alpha_a + \alpha_r \quad (2.3)$$



Rys. 1. Współczynnik ekstynkcji dla różnych stanów atmosfery

Źródło: Wyrębski W.F.: *Laserowa technika wojskowa*,
Wydawnictwo MON, Warszawa 1982.

Można wyodrębnić szereg pasm falowych, w których powietrze przepuszcza promienie światłne znacznie lepiej niż w pozostałych odcinkach widma optycznego. Są to np. pasma: $0,5 \div 0,9 \mu\text{m}$; $1,0 \div 1,1 \mu\text{m}$; $1,2 \div 1,3 \mu\text{m}$; $1,55 \div 1,75 \mu\text{m}$; $2,1 \div 2,4 \mu\text{m}$; $3,4 \div 4,1 \mu\text{m}$ oraz $8 \div 12 \mu\text{m}$. Wybitnie ujemny wpływ na rozchodzenie się wiązki laserowej w atmosferze mają opady śniegu lub deszczu oraz tzw. aerozole, czyli zawiesiny w powietrzu bardzo

drobnych ciał ciekłych (mgła, chmury) lub stałych (dym, pył, kryształki lodu itp.). Na rysunku 1 przedstawiono wpływ tłumienia atmosfery dla różnych stanów pogody i niektórych długości fal.

Drugim elementem tłumienia wiązki laserowej w atmosferze jest zjawisko rozpraszania światła. Mamy tu do czynienia z oddziaływaniem fali świetlnej (czyli fotonów) z cząsteczkami gazów zawartych w atmosferze. Intensywność tego procesu zależy w znacznym stopniu od długości fali zgodnie z prawem Rayleigha [3]. Prawo to wyraża zależność natężenia promieniowania od drogi przebytej przez wiązkę świetlną. Tłumienie promieniowania, zachodzące w wyniku dyspersji relejowskiej, może być w widzialnej części widma znacznie silniejsze aniżeli absorpcja molekularna.

Rozchodzenie się promieni lasera w wodzie podlega podobnym prawom jak i podczas propagacji w atmosferze. Monochromatyczna wiązka światła laserowego wnikając w głębi toni morskiej ulega osłabieniu podobnie jak i w powietrzu:

$$I = I_0 \exp(-\epsilon h) \quad (2.5)$$

gdzie: ϵ – współczynnik osłabienia światła monochromatycznego o długości λ ;
 h – długość drogi jaką przebywa wiązka światła lasera w wodzie.

Analogicznie jak w powietrzu na tłumienie wnoszone przez ośrodek wodny składają się – pochłanianie energii wiązki i jej rozproszenie. Wartość tłumienia światła w wodzie jest oczywiście znacznie większa niż w atmosferze. Gdy w wodzie znajdują się rozpuszczone sole i zawiesiny – zwiększają w istotny sposób zarówno pochłanianie, jak i rozproszenie. Oba te czynniki przyczyniają się do skrócenia zasięgu rozprzestrzeniania się promieni lasera. Wpływ absorpcji i rozproszenia zależą od przezroczystości wody. W wodzie o dużej przezroczystości najsilniej występuje proces absorpcji, a w wodzie mętnej zasadniczą rolę w tłumieniu odgrywa rozproszenie.

Na rys. 2 przedstawiono charakter zmian współczynnika absorpcji światła w zależności od długości fali w trzech różnych ośrodkach: woda destylowana, Ocean Atlantycki i Morze Bałtyckie. Często charakterystykę stanu hydrooptycznego wody morskiej wprowadza się wielkość $\delta = \epsilon^{-1}$ zwaną długością osłabienia. Długość osłabienia oznacza grubość warstwy wody powodującą e-krotne osłabienie energii przenikającej ją światła laserowego. Znajomość współczynnika osłabienia ϵ w określonym obszarze morza zapewnia określenie głębokości zanurzenia okrętu do której możliwy będzie odbiór sygnałów z wykorzystaniem łącza laserowego. Aby dokładnie wyznaczyć tą głębokość należy uwzględnić również parametry techniczne charakteryzujące układ nadawczy i odbiorczy, takie jak moc promieniowania lasera oraz czułość detektora.

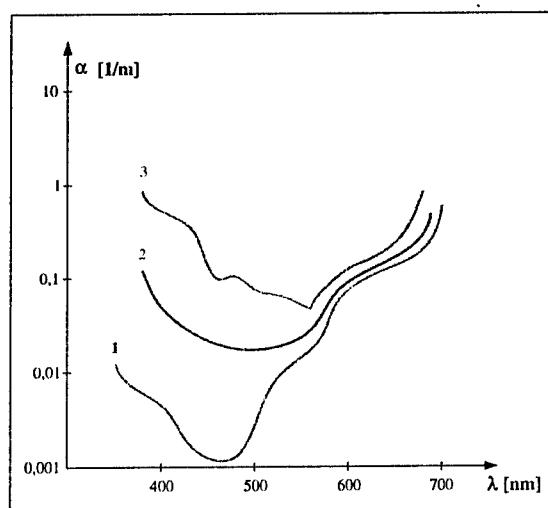
Z badań eksperymentalnych wynika, że w wodzie najsłabsze tłumienie występuje dla światła lasera o długości [4] w przedziale 480 nm do 550 nm. W stosunkowo czystych wodach oceanicznych optymalna długość światła przypada na niebiesko-zielony obszar widma. W wodach morskich zanieczyszczonych, zwłaszcza składnikami organicznymi należy stosować lasery pracujące na nieco dłuższych falach (w Morzu Bałtyckim 580 nm).

Tablica 2.1

Tłumienie wiązki światła lasera na drodze 100 m

Długość światła λ [nm]	Długość osłabienia δ [m]	Współczynnik osłabienia α [1/m]	Tłumienia na drodze 100 m
465	13,7	0,073	31,7dB
500	11,6	0,086	37,3dB
530	9,1	0,11	48,2dB

Źródło: T.G.Palopoulos, „Blue-Green Dye Lasers for Underwater ILLumination”, Naval Engineers Journal, FALL 2002, 31-40.



Rys.2. Zależność współczynnika absorpcji od długości fali w:
1 – wodzie destylowanej; 2 – Atlantyku; 3 – Morzu Bałtyckim

Należy przy tym dodać, że w określonym przypadku należy szczegółowo rozpatrzeć które fale z zakresu niebiesko-zielonego należy użyć. Przeprowadzono badania dwóch systemów laserowych – jeden pracujący z falą o długości 465 nm i drugi o długości 530 nm. Wyznaczono [5] poziom sygnału w odległości 100 m od źródła. Stwierdzono, iż tłumienie wiązki światła laserowego o długości 530 nm jest o 16,5 dB większe niż w przypadku fali o długości 465 nm.

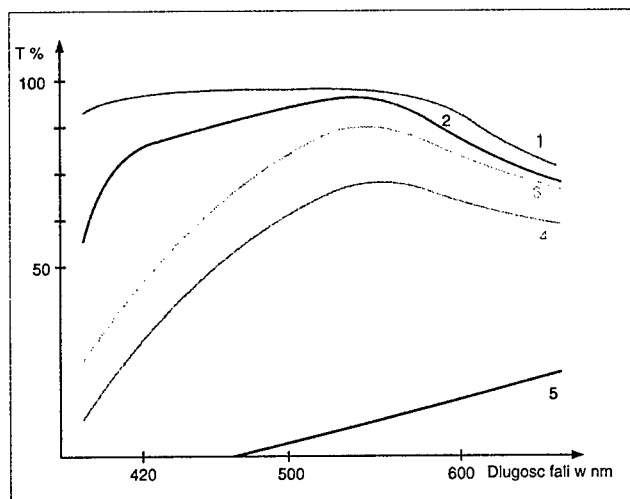
Charakterystykę warunków propagacji promieni światła określa się za pomocą transmisji (przezroczystości) wody, którą wyznacza się jako wartość wyrażającą stosunek energii światła przepuszczanego przez warstwę wody o grubości $h = 1$ m, do energii fali wnikażącej do tej warstwy. Przezroczystość wody nie jest wielkością stałą, a zależy od długości fali zastosowanego światła laserowego. W tablicy 2.2 zaprezentowano [6] wartości współczynnika osłabienia, długości osłabienia, przezroczystości i głębokości wnikania. Dane te przedstawione są dla różnych długości fali promieni laserowych rozchodzących się w wodzie destylowanej.

Tablica 2.2

Właściwości propagacji promieni laserowych wytwarzanych przez różne lasery

Rodzaj lasera	λ [nm]	T [%]	ϵ [1/m]	δ [m]	$Z=7\delta$ [m]
Laser neodymowy typu YAG	1064,8	0,001	11,000	0,09	0,6
Laser rubinowy	694,3	58,9	0,537	1,8	12
Laser He-Ne	632,8	78,1	0,243	2,1	14
Laser z drugą harmoniczną YAG	523,4	95,2	0,049	20,4	140
Laser argonowy	488,0	97,2	0,027	37,0	260
Laser kadmowy	411,6	96,2	0,039	25,7	175
Druga harmoniczna lasera rubinowego	347,7	95,4	0,047	21,3	150

Z tablicy 2.2 widać, jak istotne jest dostosowanie długości fali lasera do pasma optymalnej transmisji światła w wodzie. Przezroczystość wody jest wielkością charakterystyczną dla poszczególnych zbiorników morskich. Podlega ona, zwłaszcza w małych zamkniętych morzach, wyraźnym zmianom sezonowym. Na rys. 3 przedstawiono charakterystyki przezroczystości wybranych wód.



Rys. 3. Charakterystyka przezroczystości wybranych wód:

1 – wody destylowanej, 2 – Północnego Atlantyku, 3 – Morza Północnego,
4 – Morza Bałtyckiego, 5 – wody bagiennej

Naturalnym sposobem, wybranym przez konstruktorów, do zwiększenia głębokości wnikania promieni laserowych jest zwiększanie mocy lasera. Jednak występują ograniczenia nie tylko ze względów technicznych. W rzeczywistych warunkach w morzu nie można gęstości mocy światła laserowego zwiększać do dowolnej wielkości. Przy gęstości mocy światła laserowego rzędu 10^{13} W/cm² występują wzdłuż wiązki laserowej w wodzie morskiej gwałtowne jonizacje drobin i intensywne wrzenie wody, wykluczające w tym stanie możliwość zastosowania lasera do łączności.

3. SATELITARNE SYSTEMY ŁĄCZNOŚCI Z ZASTOSOWANIEM LASERÓW

Od 1984 roku prowadzono eksperymenty z systemem łączności, który umożliwi transmisję informacji w relacji „samolot- zanurzony okręt podwodny”.

Laser cieczowy (w tym barwnikowy) z lampą pompującą ma możliwość ciągłej zmiany długości fali w zakresie widzialnym. Z tego względu jego wykorzystanie zapewnia dobranie na drodze eksperymentu optymalnej długości fali, tak aby występowało minimalne pochłanianie w wodzie morskiej. W związku z tym, że większość laserów tego typu charakteryzuje się niewielką efektywnością i krótkim czasem pracy obecnie są one zastępowane laserami wykonanymi na bromku rtęci i przez lasery gazowe wykonane na chlorku ksenonu. Te ostatnie mają możliwość przestrajania, aż do uzyskania widma optycznego w kolorze niebiesko-zielonym.

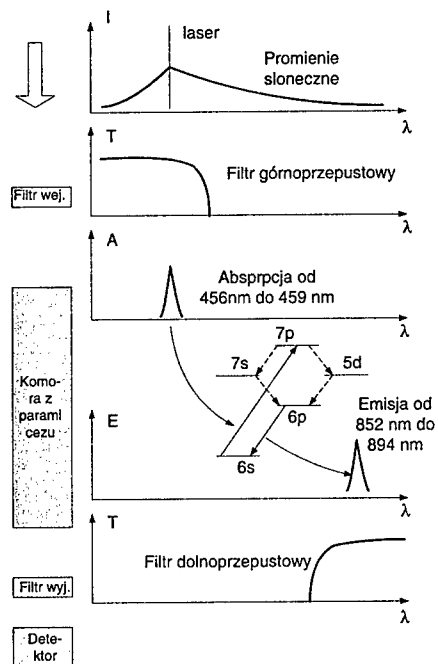
Badania wykazały, że laser gazowy charakteryzuje się największą mocą promieniowania i największą sprawnością w zakresie widma nadfioletowego. Firmie „Nortrop” podczas realizacji projektu systemu łączności laserowej udało się opracować laser gazowe na chlorku ksenonu ze wzbudzeniem za pomocą pola elektrycznego. Uzyskano wiązkę lasera o promieniu w zakresie fal optycznych o długości 308 nm.

Nawet w zakresie długości fal optycznych [4] odpowiadającej „*Jerlov Minimum*” (dla długości fali $\lambda = 450$ nm), tłumienie wody morskiej nie jest do pominięcia, a sygnały docierające do zanurzonego okrętu podwodnego mogą być dość słabe. W związku z tym odbiorniki znajdujące się na okrętach podwodnych powinny mieć bardzo dużą czułość. Wymaganie dużej czułości wprowadza dodatkowe komplikacje - światło słoneczne zawiera znaczną ilość składowych z przedziału niebiesko-zielonego, które mogą również wnikać w wody oceanu i wprowadzać zakłócenia do odbieranego sygnału. Jedyną drogą rozwiązania tego problemu jest wykorzystanie zjawiska wynikającego z różnicy między bardzo wąskim widmem promieniowania lasera niebiesko-zielonego i dużo szerszym widmem promieni słonecznych.

W nowszych rozwiązaniach systemów łączności laserowej jeszcze przed wejściem odbiornika instaluje się specjalne filtry eliminujące wpływ promieni słonecznych. Według opinii naukowców amerykańskich zastosowanie atomowych filtrów rezonansowych pozwoli zwiększyć maksymalną głębokość wnikania promienia lasera do 100 m lub 15 razy skrócić czas nadawania przy jednoczesnym zmniejszeniu mocy promieniowania nadajnika. Filtr ten charakteryzuje bardzo wąskie pasmo przepuszczania - obejmujące zakres przedziału niebiesko-zielonego, tak aby przepuścić fotony promienia lasera do detektora, a jednocześnie uniemożliwić przedostanie się fotonów promieni słonecznych. W dodatku ten wąskopasmowy filtr musi mieć szerokie pole obserwacji. Jak z tego wynika mamy do czynienia z dwoma przeciwstawnymi wymaganiami, które trudno [7] spełnić. Najbardziej zbliżony do potrzeb zastosowania w układzie odbiorczym na okręcie podwodnym jest „atomowy filtr rezonansowy” lub „ARF” zwany również jako ograniczony kwantowo detektor z rezonansem optycznym (ang. *QLORD - quantum limited optical resonance detector*). Stosunek szerokości kanału optycznego $\Delta\lambda$ do środkowej częstotliwości (długości fali λ_0) pasma w którym te filtry mogą być stosowane wynosi $\Delta\lambda/\lambda_0 \approx 10^{-6}$. Tak więc, dla środkowej częstotliwości tj. dla fali o długości 500 nm, powinno być wąskie pasmo o rozbieżności [3] wiązki $\sim 0,005\text{\AA}$. Filtr ARF zbudowany na bazie par cezu jest szczególnie przydatny w zastosowaniu do układów w systemie łączności z okrętami podwodnymi.

Działanie układu filtracji można zilustrować na podstawie rys. 4. Filtr wejściowy przepuszcza tylko światło barwy niebiesko-zielonej, które następnie dostaje się do komory z parami cezu. W parach cezu następuje absorpcja kwantów promieniowania światła, o dłu-

gości fali 456nm lub 459nm, czemu odpowiada przejście z poziomu energetycznego 6s na poziom 7p. Kolejno następuje bezpromieniste przejście do poziomu 6p, poprzez poziomy 7s lub 5d. Wówczas gdy, nastąpi powtórne rozładowanie do obudowy wystąpi wyemitowanie fotonów w paśmie podczerwieni (w zakresie 852 nm do 894 nm). Wyjściowy filtr przepuszcza tylko promieniowanie, z zakresu podczerwieni, oddziaływujące na detektor. Tak więc pasma dwóch filtrów optycznych (wej. i wyj.) nie zachodzą na siebie i światło słoneczne nie powinno docierać do detektora. Fotony światła lasera z zakresu niebiesko-zielonego mogą oddziaływać na detektor gdyż w komorze z parami cezu następuje konwersja widna w wyniku kolejnych procesów absorpcji i reemisji. Filtr ARF ma bardzo wąskie pasmo wymagane dla odfiltrowania zbędnych składowych z promieniowania słonecznego a jednocześnie zapewnia odbiór promieni lasera z zakresu widma niebiesko-zielonego.



Rys.4. Zasada działania filtru ARF

gdzie: I – intensywność światła słonecznego; T – tłumienie filtru;

A – absorpcja w komorze z parami cezu; E – emisja w komorze cezu.

Te same wymagania, które odnoszą się do filtru cezu stosuje się również w przypadku lasera niebiesko-zielonego znajdującego się na satelicie. Laser powinien być dokładnie strojony odpowiednio do częstości drgań elektromagnetycznych odpowiadających przejściu z poziomu 6s do 7p; długość fali światła o długości 456nm lub 459 nm powinna [8] mieścić się w oknie o szerokości ok. 0,01Å. Wymagana moc promieniowania takiego lasera jest na poziomie kilowat. [9].

4. PODSUMOWANIE

Woda morska jest specyficznym środowiskiem rozchodzenia się fale elektromagnetycznych, akustycznych i optycznych. Badania wykazały, że najmniejsze straty występują w przypadku rozprzestrzeniania się fal elektromagnetycznych w zakresach LF, VLF

i ELF. W przekazywaniu informacji znalazły zastosowanie również urządzenia laserowe. Przy tym fale radiowe o bardzo niskiej częstotliwości przenikają nawet do głębokości 100-150 m poniżej powierzchni morza, natomiast promienie lasera niebiesko-zielone nawet do 300 m.

Eksperymentalne systemy satelitarnej łączności laserowej przewyższają swoimi możliwościami systemy wykorzystujące fale elektromagnetyczne przede [10] wszystkim ze względu na: odporność na zakłócenia; niezawodność; szybkość transmisji; a nade wszystko ze względu na głębokość wnikania fali optycznej w toń morza. Niekorzystne zjawiska występujące na drodze od satelity do zanurzonego okrętu podwodnego mają wpływ na wymagania energetyczne w stosunku do nadajnika laserowego. Z obecnie znanych laserów najodpowiedniejszym do zastosowania są te które pracują z falą o długości 530 nm i mocą minimum 1W. Od lat prowadzone są prace nad zastosowaniem laserów niebiesko-zielonych różnych typów: laser na bromku rtęci, lasery argonowe, lasery na parach miedzi itd.

W ostatnich latach w wielu instytutach [11] badawczych prowadzi się szeroko zakrojone badania nad utworzeniem systemów łączności laserowej w zastosowaniach telemetrycznych. Często są to złożone systemy w których częścią kanału teletransmisyjnego jest łącze laserowe, na odcinku statek powietrzny – toń morza. Innym zastosowaniem militarnym są systemy [12] wykrywania obiektów znajdujących się pod powierzchnią wody.

BIBLIOGRAFIA

- [1] Kosiński A.: *Podwodna łączność laserowa*, mat. z konferencji KKRRiT Wrocław 2003, s. 437-440.
- [2] *Naval satellite communications: anticipating the next generation*, Maritime Defence, vol 21, no.1, January/February 1996.
- [3] W. P. Risk, T.R. Gosnell, A.V. Nurmikko: *Compact blue-green laser*, Cambridge University Press 2003
- [4] Theodore G. Pavlopoulos: *Blue-Green Dye Lasers for Underwater Illumination*, Naval Engineers Journal, FALL 2002, 31-40.
- [5] Honey R.C.: „*Laser in the Underwater Environment*” Opt. Spectra, April 1972.
- [6] H.Z. Wrembel.: *Wybrane problemy optyki morza*, Zeszyty naukowe WSMW, Nr , Gdynia 1974. s. 81-97.
- [7] Gelbwachs J.A.: „*Atomics resonance filters*”. JEEE 1988, J. Quant. Electron.,24, 1266-1277.
- [8] Lesile K.R.: „*Alexandrine solid state blue laser*”, Proc. SPIE 1995, 2380, 82-87.
- [9] Laser Focus World: „*Blue-green laser links to subsystem*”, April 1980, 14-18.
- [10] Ferguson G.D.: „*Blue-Green Lasers for Underwater Applications*” SPIE 64, 1974.
- [11] Raport by: B. Mooney, D. Wals, June 1994, http://www.wtec.org/loyola/subsea/c8_s3.htm
- [12] Burnham R., E.J. Schimitschek: „*High Power Blue Green Lasers*”, Laser Focus World, June 1981.

BLUE-GREEN OF LASER RAYS IN USE TO CONTACT WITH SUBMEGED OBJECTS

Abstract

This article new directions of investigations and achievements were presented in range of sending of information to plunged submarines with use of lasers' rays. Methods of broadcast of information were introduced to submarines with utilisation of rays in the range of blue-green laser radiation spectrum. The blue-green laser radiation is capable of penetration to depths of more than 90 m. To obtain the widest possible coverage it was decided to use satellites as relays between ground stations and submarines.

Andrzej Kosiński, Bogusław Pilek

Zakład Łączności, Akademia Marynarki Wojennej

MONITORING RADIOWY OBIEKTÓW SZYBKO PRZEMIESZCZAJĄCYCH SIĘ NAD MORZEM

Streszczenie

W artykule przedstawiono koncepcję wykorzystania systemu „WOJOWNIK” do zdalnego monitoringu obiektów latających, a w szczególności samolotów i rakiet. Opisano budowę i działanie systemu. Zaprezentowano koncepcje poszerzenia możliwości zastosowania urządzenia. W końcowej części zamieszczono uwagi dotyczące badań poligonowych urządzenia monitorującego.

1. WSTĘP

System „WOJOWNIK” przeznaczony jest do automatycznego śledzenia położenia i parametrów technicznych obiektów latających w czasie rzeczywistym. System opracowano i wykonano w Akademii Marynarki Wojennej jako odpowiedź na potrzebę zdalnej diagnostyki technicznej rakiet typu woda-woda. Jednocześnie zdolność systemu do określania położenia monitorowanego obiektu w danej chwili, pozwoliła zapewnić bezpieczeństwo ćwiczeń morskich.

Pierwsze próby systemu przeprowadzono w 1993r na poligonie morskim monitorując lot i parametry przeciwokrętowych kierowanych pocisków raketowych. Po udanych próbach śledzenia lotu rakiet, i niezbędnej modyfikacji systemu, przeprowadzono udane testy systemu na samolotach i śmigłowcach Lotnictwa Marynarki Wojennej.

2. BUDOWA I OPIS DZIAŁANIA SYSTEMU „WOJOWNIK”

System „WOJOWNIK” składa się z urządzenia montowanego na obiekcie latającym (także okręcie lub pojeździe) i stanowiska dyspozytorskiego zainstalowanego na brzegu. Na obiekcie latającym obróbką informacji z Globalnego Systemu Pozycyjnego (GPS), przetwarzaniem wybranych parametrów diagnostycznych oraz cyfrową transmisją radiową realizuje specjalizowany sterownik wykonany w oparciu o procesor jednoukładowy. Na stanowisku dyspozytorskim do obróbki informacji, rejestracji i wizualizacji trajektorii lotu na mapie elektronicznej oraz parametrów technicznych wykorzystuje się radiowy zestaw odbiorczy i standardowy komputer.

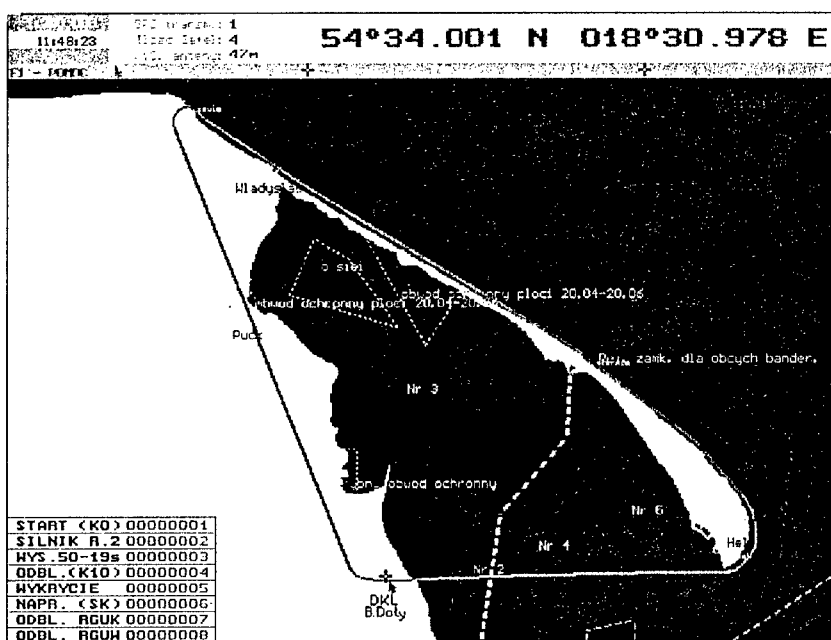
Urządzenie „WÓJOWNIK” wykonane jest w technologii montażu powierzchniowego, a zastosowane podzespoły spełniają wysokie wymagania klimatyczno-elektryczne, zapewnia to dużą niezawodność pracy systemu. Poszczególne układy elektroniczne umieszczone są w wyfrezowanych komorach jednobryłowej obudowy wykonanej z aluminium, stanowiącej dodatkowo radiator odprowadzający ciepło. Zapewnia to dużą wytrzymałość mechaniczną oraz ekranowanie poszczególnych układów.

Parametry techniczne urządzenia „WOJOWNIK”

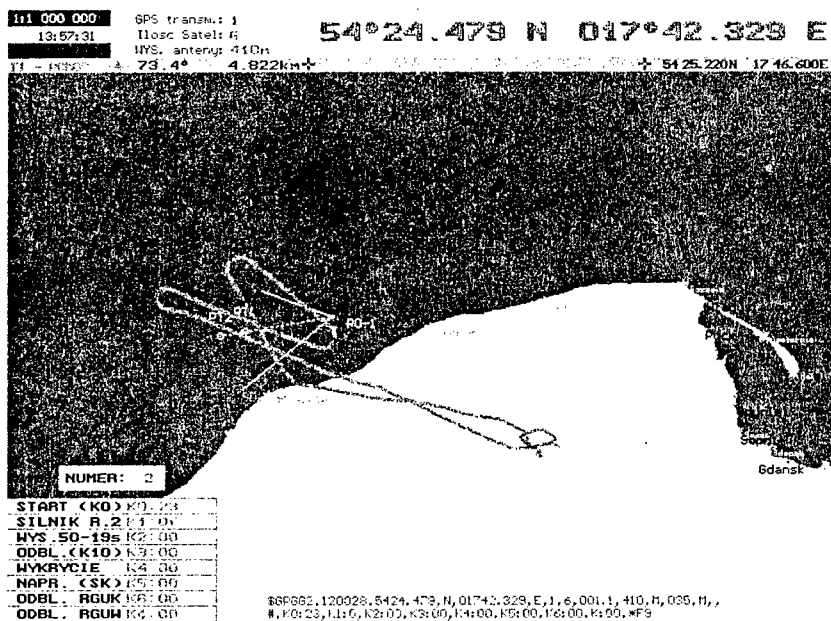
- napięcie zasilania $24 \div 35$ V lub opcja 12 V;
- pobór mocy ze źródła zasilania max 150 W ;
- moc prom. nadajnika w.cz. 30 W lub opcja 50 W;
- częstotaś odświeżania danych diagnostycznych 1 s;
- częstotliwość radiowa sygnału z zakresu UHF;
- ciężar urządzenia na obiekcie latającym 2,7 kg;
- wymiary 195 x 90 x 235.

Urządzenie zapewnia transmisję danych z monitorowanego obiektu do stanowiska dyspozytora na brzegu (lądzie). Monitorowany obiekt może poruszać się z dużą prędkością (max 500 m/s) i to narzuca określone krytyczne wymagania na parametry techniczne systemu, jak i na zastosowaną technologię. Dotyczy to całości urządzenia oraz poszczególnych podzespołów i układów elektronicznych. Urządzenie musi być odporne na duże przeciążenia, zmianę temperatury otoczenia i wibracje. Ponadto ze względu na przeznaczenie danych pozyskiwanych z monitorowanego obiektu (rakiety) transmisja odbywa się w czasie rzeczywistym. Z tym również ma związek wymagania na szybkość radiowej transmisji danych, która powinna zapewnić zdolność do przekazania niezbędnych parametrów obiektu przy odpowiednio niskiej stopie błędów. Dlatego po próbach systemu i uwzględnieniu poziomu zakłóceń radiowych na poligonie morskim podczas ćwiczeń, przyjęto szybkość transmisji danych 2400 bit/s. Ponadto system charakteryzują dwa ważne parametry: dokładność określenia współrzędnych monitorowanego obiektu, która wynosi ok. 20 m i maksymalny zasięg monitoringu do 100 km.

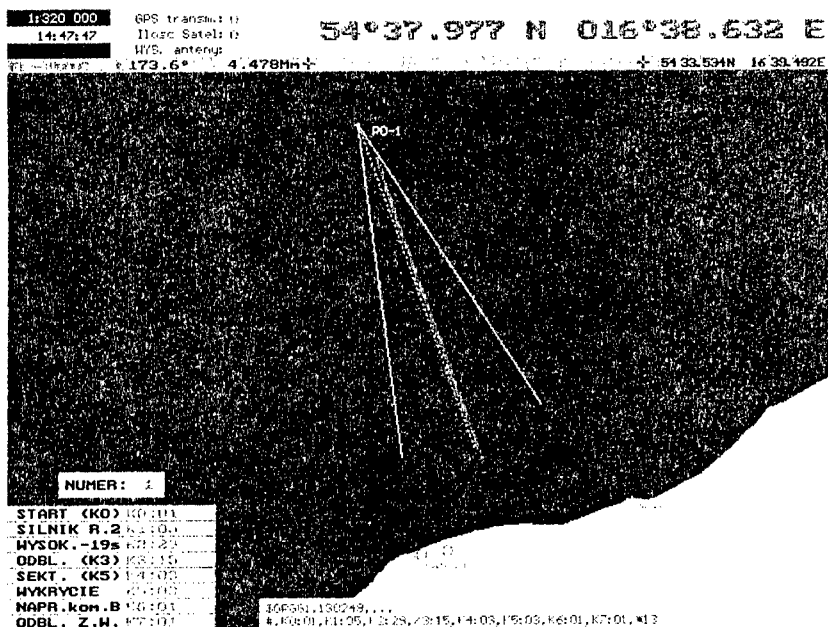
Dotychczas przeprowadzono udane próby urządzenia „WOJOWNIK” na samolotach TS 11 ISKRA (rys. 1, 2), śmigłowcach oraz rakietach typu woda – woda (rys. 3). Próby wykazały jak wysokie wymagania niezawodnościowe muszą spełniać tego typu urządzenia. Przede wszystkim niezbędna jest technologia montażu powierzchniowego poszczególnych układów elektronicznych. W przypadku samolotów TS11-Iskra oraz śmigłowców Sokół prowadzono zdalny monitoring położenia, natomiast ze względów formalnych (konieczność przeprowadzenia wymaganych prób kwalifikacyjnych) nie podłączono wejść przetworników analogowo-cyfrowych urządzenia do wskaźników oraz czujników samolotu.



Rys.1. Zarejestrowana trajektoria lotu samolotu TS-11 Iskra w rejonie zatoki Gdańskiej.



Rys. 2. Zarejestrowana trajektoria lotu samolotu TS-11 Iskra w rejonie (Śemirowice-Ustka).



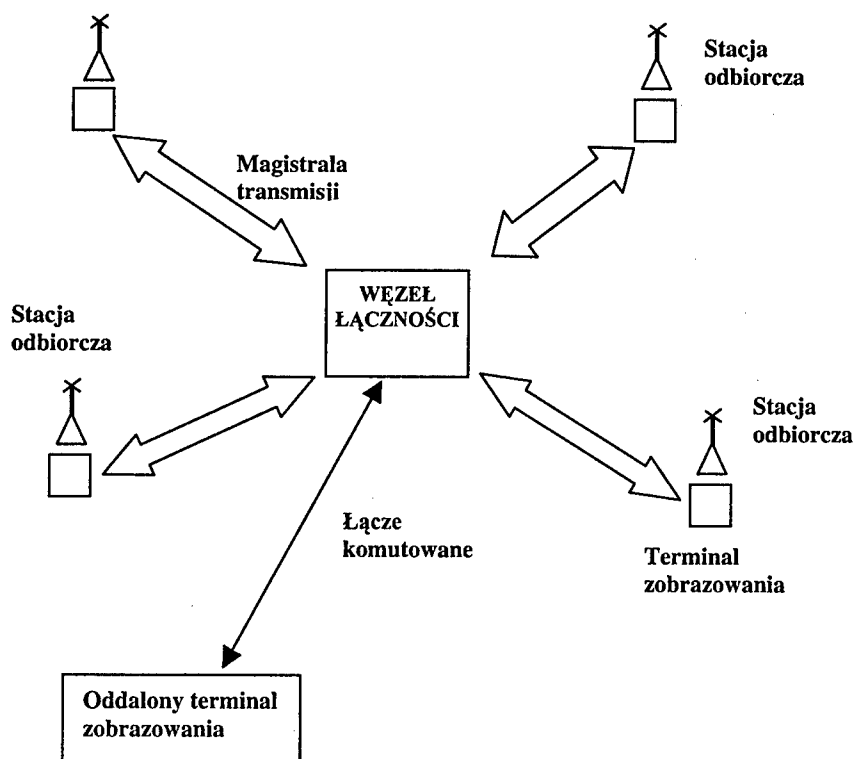
Rys. 3. Zarejestrowana trajektoria lotu i parametrów rakiety 4k31.

3. ZASTOSOWANIE URZĄDZENIA „WOJOWNIK” DO MONITORINGU STATKÓW POWIETRZNYCH W LOCIE

Urządzenie WOJOWNIK-M może być zastosowane w obiektach ruchomych (latających) posiadających źródło zasilania z zakresu 24-35 V lub 12 V oraz prędkością nie przekraczającą 500 m/s. (parametr zastosowanego modułu GPS). Monitorwane parametry techniczne (8 wartości analogowych) mogą mieć rozdzielczość 10 bit. (1024 wartości) co wyklucza monitorowanie bardzo precyzyjnych parametrów. Częstotliwość odświeżania wartości parametrów 1Hz umożliwia monitorowanie wolnozmiennych procesów zachodzących w obiekcie latającym, np. temperatura silnika, poziom paliwa, aktualna prędkość, wysokość itp. Urządzenie przystosowane jest do pracy ciągłej w temperaturze od -20 do +50 °C. Jedno bryłowa obudowa-radiator zapewnia dużą wytrzymałość konstrukcji i odporność na zewnętrzne czynniki mechaniczne. W przypadku monitoringu samolotów lotnictwa morskiego RP możliwe jest śledzenie następujących parametrów:

- Współrzędne położenia;
- Wysokość;
- Prędkość;
- Poziom paliwa;
- Temperatura silnika;
- Temperatura otoczenia;
- Stan uzbrojenia;
- Dane z zainstalowanych przyrządów pomiarowych.

Przy modyfikacjach programowych sterownika odpowiedzialnego za wyznaczenie szczeliny czasowej transmisji dla danego obiektu, można przy dodatkowym zwielokrotnieniu częstotliwości, zapewnić jednoczesną obserwację do kilkunastu obiektów w danej strefie podlegającej monitorowaniu. Obszar strefy ograniczony jest zasięgiem kanału łączności radiowej pomiędzy nadajnikiem urządzenia „WOJOWNIK” a odbiorczą stacją bazową. Aby zwiększyć zasięg działania systemu, tak aby obejmował większy obszar, np. obszar kraju, należałoby umieścić stacje bazowe (odbiorcze) systemu „WOJOWNIK” w odstępach nie większych niż 80 km. Aby zapewnić zobrazowanie położenia i parametrów technicznych samolotów w całej strefie zainteresowania Marynarki Wojennej, a przede wszystkim w centrach kierowania lotów, można wykorzystać komutowane i niekomutowane łącza transmisji danych lub zastosować „linki” radioliniowe pomiędzy tymi ośrodkami. Schemat łączności (transmisji danych) pomiędzy poszczególnymi stacjami bazowymi oraz ośrodkami zobrazowania danych przedstawia rys.4.



Rys. 4. Schemat łączności (transmisji danych) pomiędzy poszczególnymi stacjami bazowymi oraz ośrodkami zobrazowania danych.

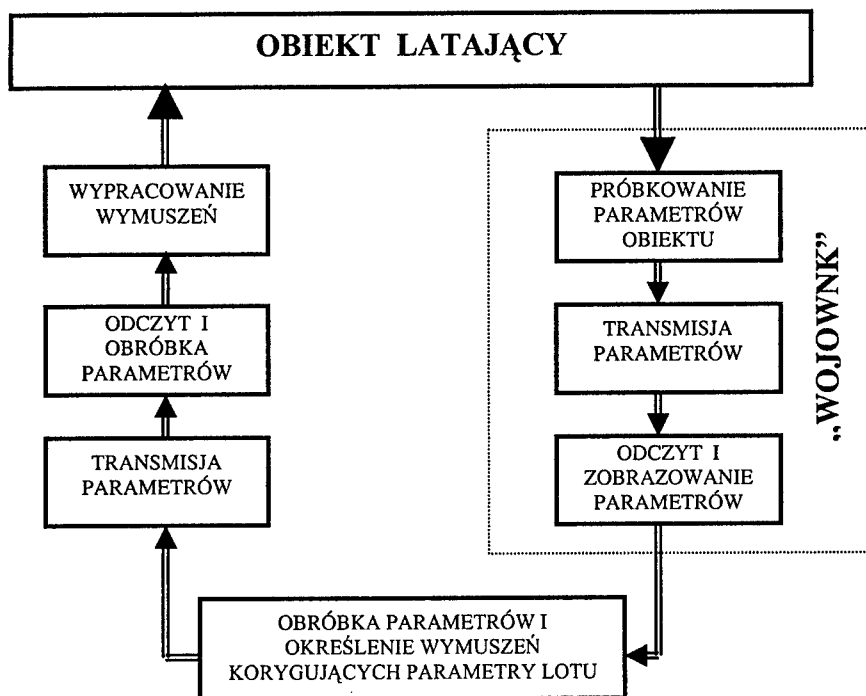
Aby system monitorowania samolotów lotnictwa MW był w pełni dostępny przez zainteresowane ośrodki, należałoby stworzyć sieć transmisji danych. Użytkownik wówczas mógłby w dowolnym punkcie sieci podłączyć swój terminal odbiorczy i obserwować sytuację w całym obszarze działania systemu. Oddaleni użytkownicy systemu mogliby korzystać z komutowanych łączy transmisji danych, co umożliwiłoby podłączenie terminalu zobrazowującego w dowolnym punkcie kraju. Terminal odbiorczy podłączony w dowolnym punkcie sieci transmisji danych powinien umożliwić użytkownikowi lokalizację na mapie elektronicznej samolotu będącego w powietrzu w dowolnym punkcie obszaru działania systemu wraz z jego aktualnymi parametrami technicznymi. Dodatkowe funkcje, jakie mógłby spełniać system to:

- automatyczna rejestracja trasy lotu wraz z parametrami technicznymi;
- automatyczne alarmowanie w przypadku zaniku sygnału lub zmiany ustalonych wartości granicznych parametrów np.: wysokość, prędkość, poziom paliwa itp.;
- automatyczne przesłanie sygnału zagrożenia lub awarii z obiektu latającego;
- symulacja maksymalnego zasięgu lub czasu lotu dowolnie wybranego obiektu.

4. STEROWANIE OBIEKTAMI LATAJĄCYMI Z WYKORZYSTANIEM URZĄDZENIA WOJOWNIK

Sterowanie obiektów latających może być realizowane w zakresie parametrów monitorowanych przez urządzenie „WOJOWNIK” oraz może wykraczać poza ten zakres, przy czym tylko w pierwszym przypadku istnieje pełny obraz stanu obiektu latającego oraz realizacji wymuszenia sterującego. W przypadku wykorzystania tego urządzenia do sterowania obiektem latającym stanowi on element sprzężenia zwrotnego zobrazowującego efekt sterowania. Algorytm sterowania parametrami obiektu latającego z wykorzystaniem urządzenia „WOJOWNIK” przedstawia rys.5.

Jak już wcześniej wspomniano badane urządzenie zdolne jest do monitorowania ośmiu parametrów technicznych obiektu latającego (sygnały analogowe) z rozdzielczością 10 bit oraz jego położenia z dokładnością do 20 m. Istotnym czynnikiem określającym zakres sterowania parametrami oraz rodzaj obiektu i jego dynamikę, jest odstęp czasu pomiędzy kolejnymi uaktualnieniami wartości parametrów. W przypadku urządzenia „WOJOWNIK” jest to czas 1 sek., który wraz z dokładnością określenia współrzędnych istotnie ogranicza zakres sterowania lotem obiektu. I tak dla samolotu TS-11 Iskra, który może poruszać się z prędkością 700 km/h, od momentu pomiaru położenia do kolejnego wymuszenia kierunku lotu samolot może przebyć drogę do 400 m. W przypadku samolotów bezzałogowych bez systemów autopilota sterowanie zdalne z wykorzystaniem urządzenia „WOJOWNIK” byłoby mocno ograniczone. Duże znaczenie ma zastosowanie sterowania dla rakiet z układem samonaprowadzania typu (woda-woda), które posiadają żyroskopowy system kontroli parametrów lotu.



Rys. 5. Algorytm sterowania obiektem latającym

5. PODSUMOWANIE

Podczas zdalnego monitoringu lotu obiektów szybko poruszających się nad morzem napotkano na szereg problemów natury techniczno-organizacyjnej. Problemy te miały ścisły związek ze specyfiką warunków pracy urządzeń zdalnego monitoringu, a zwłaszcza dużą prędkością obserwowanych obiektów: od 50m/s dla śmigłowców i samolotów, do 400m/s w przypadku zdalnego monitorowania lotu i parametrów, przeciwookrętowych kierowanych pocisków rakietowych (PoKPR). Te specyficzne warunki wymusiły zastosowanie w urządzeniu monitorującym jedno-bryłowej obudowy wraz z radiatorem, która jednocześnie zapewniała chłodzenie tranzystorów wzmacniacza mocy w.cz. Ze względu na duże przeciążenia i udary na jakie narażone jest urządzenie, zwłaszcza podczas startu rakiety, poszczególne moduły urządzenia muszą być wykonane w technologii montażu powierzchniowego (SMD) i umieszczone w wyfrezowanych komorach wypełnianych substancją dielektryczną. Równie istotny wpływ na pracę systemu miał morski klimat pracy urządzeń elektronicznych. Ze względu na dużą wilgotność i zasolenie powietrza płytki elektroniczne muszą być lakierowane i całość urządzenia umieszczona jest w hermetycznej obudowie. Ogromny wpływ na niezawodność pracy systemu monitorującego miał tzw. „czynniki ludzki” (błędy operatorów i niewłaściwe użytkowanie systemu).

Tablica 1

Statystyka niesprawności systemu

WYKAZ PRZYPADKÓW NIEWŁAŚCIWEJ PRACY SYSTEMU „WOJOWNIK” PODCZAS BADAŃ NA POLIGONIE MORSKIM (Liczba monitorowanych obiektów 98)				
Typ awarii	Lp.	Przyczyna niesprawności	Liczba przypadków	Uwagi
Awarie związane z pracą GPS	1.	Brak synchronizacji urządzeń GPS przed startem rakiety (nieznana przyczyna)	4	Synchronizacja w czasie lotu
	2.	Niewłaściwe podłączenie sprzęgu anteny relingowej GPS z anteną na pokrywie rakiety	6	Błąd operatora (synchronizacja w czasie lotu)
	3.	Utrata synchronizacji GPS podczas lotu rakiety	5	Chwilowy brak transmisji danych o położeniu rakiety
Awarie związane z radiową transmisją danych i sprawnością techniczną systemu.	1.	Zakłócenia radiowej transmisji danych z rakiety.	4	Przypadki w których stopa błędów przekraczała 5%
	2.	Uszkodzenie urządzenia nadawczego podczas startu rakiety	3	Brak nośnej sygnału w.cz.
	3.	Uszkodzenie urządzenia nadawczego podczas lotu	2	Brak nośnej sygnału w.cz.
	4.	Awaria modemu	1	
	5.	Brak transmisji parametrów technicznych rakiety przy prawidłowej transmisji położenia	1	Nieprawidłowe podłączenie przetworników A/C przez operatora
	6.	Awaria stanowiska odbiorczego systemu	1	Zawieszenie się systemu operacyjnego komputera

Z powyższych danych wynika, że tzw. „czynniki ludzkie” (obsługa operatorska) stanowi bardzo istotny problem dla zachowania właściwej niezawodności systemu. Na 27 przypadków częściowej lub pełnej niesprawności w 9 przypadkach przyczyną był błąd operatora. Drugim newralgicznym punktem pracy systemu jest moduł GPS. Przed wystreleniem obiektu należy zsynchronizować moduł GPS. Natomiast nie stwierdzono istotnego wpływu efektu Dopplera na stopę błędów radiowej transmisji danych w paśmie UHF dla modulacji częstotliwościowej (sygnałem AFSK). Próby realizowano w dwóch różnych wariantach mogących mieć wpływ na efekt Dopplera: obiekt latający przemieszczał się z dużą prędkością równolegle do brzegu, w drugim przypadku obiekt przemieszczał się w kierunku prostopadłym do brzegu (stanowiska odbiorczego).

RADIO-MONITORING FAST TRANSLOCATE OBJECTS ABOVE SEA LEVEL

Summary

In this article has been presented new idea used the “WOJOWNIK-M” installation to monitoring flying of the objects in peculiarity aircrafts and rockets. Described construction and operations of the systems. Has been characterised idea extension possibility application of the installation. In the latter point has been inserted notices about the research monitoring installation.

Michał Mikołajski

Katedra Systemów i Sieci Radiokomunikacyjnych, Politechnika Gdańska

BADANIA SYMULACYJNE ALGORYTMU LS-DRTMA W ZASTOSOWANIACH ANTENOWYCH

Streszczenie

W artykule przedstawiono algorytm adaptacyjny LS-DRMTA (Least Squares Despread Respread Multitarget Array), przeznaczony do kształtowania własności kierunkowych anteny inteligentnej w systemach szerokopasmowych z wielodostępem CDMA. Opisano analitycznie podstawy teoretyczne tego algorytmu. Omówiono wyniki wykonanych badań symulacyjnych i na ich podstawie scharakteryzowano praktyczną przydatność tego algorytmu.

1. WSTĘP

Rozwój algorytmów adaptacyjnych anten inteligentnych związany był dotychczas z klasycznymi metodami modulacji oraz wielodostępem. Oznaczało to, że algorytmy adaptacyjne w swojej strukturze zawierały założenie o wąskim paśmie zajmowanym przez sygnały użyteczne (patrz np. [1]). Wraz z wprowadzeniem wielodostępu opartego o technikę CDMA należało zweryfikować założenie o wąskopasmowości sygnału. Z uwagi na to, że w międzyczasie odkryto, że odbiór i nadawanie zbiorcze przynoszą także dobre efekty lepszego wykorzystania widma [2], nastąpiło znaczne spowolnienie prac badawczych nad algorytmami pracy anten inteligentnych. Zagadnienie algorytmów adaptacyjnych anteny inteligentnej dla systemów szerokopasmowych nie doczekało się własnej literatury przedmiotu, zaś prowadzone prace badawcze koncentrują się na metodach dostosowania istniejących algorytmów do pracy z sygnałami szerokopasmowymi.

Przykładem takiego podejścia jest algorytm LS-DRMTA (*Least Squares Despread Respread Multitarget Array*). Metodyka, którą zastosowano do stworzenia tego algorytmu jest bardzo zbliżona do klasycznego algorytmu LMS (*Least Mean Square*) [3]. Poważnym ograniczeniem tego algorytmu jest wymaganie, aby w macierzy antenowej występowały podukłady liczące w liczbie równej liczbie stosowanych kodów rozpraszających. Mimo tego ograniczenia algorytm warty jest prezentacji, gdyż unaocznia w jaki sposób skonstruować algorytm adaptacyjny dla sygnału szerokopasmowego. Czynnikiem pomijanym w rozważaniach jest szerokopasmowość szyku antenowego. Zakłada się a priori, że stosowany szysk w dostatecznym stopniu pokrywa zapotrzebowanie na odbiór sygnałów w zadanym paśmie. Algorytm LS-DRMTA nie jest algorytmem klasy DOA (*Direction of Arrival*) i nie

zapewnia otrzymania „wprost” estymat kierunku nadejścia sygnału. Zamiast tego wyznaczone są współczynniki wagowe szyku antenowego w celu minimalizacji określonej funkcji kosztu.

2. OPIS ANALITYCZNY ALGORYTMU

W stacji bazowej systemu CDMA sygnały rozpraszające PN poszczególnych użytkowników znane są a priori. Jeśli n -ty bit informacji wysłanej przez i -tego użytkownika został poprawnie odtworzony w odbiorniku, to pochodząca od i -tego użytkownika formacja elektromagnetyczna niosąca tę informację może zostać odtworzona za pomocą ponownego rozproszenia bitu informacyjnego przez sekwencję pseudoprzypadkową PN i -tego użytkownika. Ta formacja elektromagnetyczna może zostać zastosowana w beamformerze do wyznaczenia współczynników wagowych w sposób umożliwiający odbiór informacji od użytkownika i -tego. Zatem opóźniona wersja ponownie rozproszonego sygnału $r_i(t)$, od użytkownika i -tego, dana jest wzorem:

$$r_i(t) = \hat{b}_{in} c_i(t - \tau_i), \quad (2.1)$$

w którym:

b_{in} – bit odebrany stanowiący estymatę bitu nadanego

c_i – ciąg rozpraszający dla i -tego użytkownika

τ_i – opóźnienie dla ciągu rozpraszającego dla i -tego użytkownika.

Algorytm LS-DRMTA będzie usiłował adoptować wektor wagowy \mathbf{w} tak, aby zminimalizować funkcję kosztu. Niech $y_i(k)$ oraz $r_i(k)$ oznaczają k -tą próbkę oraz odpowiednio sygnału $y_i(t)$ oraz $r_i(t)$ – tzn. sygnału odebranego przed operacją skupiania i detekcji. Wówczas funkcja kosztu będzie opisana zależnością:

$$F(\mathbf{w}_i) = \sum_{k=1}^K |y_i(k) - r_i(k)|^2 = \sum_{k=1}^K |\mathbf{w}_i^H \mathbf{x}(k) - r_i(k)|^2, \quad (2.2)$$

w której:

K – rozmiar bloku danych równy liczbie próbek w trakcie trwania jednego bitu

$\mathbf{x}(k)$ – sygnał z zespołu antenowego próbkowany w momencie k .

Używając rozszerzenia metody Gaussa [4] możemy zapisać, że:

$$g_k(\mathbf{w}_i) = |y_i(k) - r_i(k)| = |\mathbf{w}_i^H \mathbf{x}(k) - r_i(k)|, \quad (2.3)$$

oraz

$$\mathbf{g}(\mathbf{w}_i) = \begin{bmatrix} y_i(1) - r_i(1) \\ y_i(2) - r_i(2) \\ \vdots \\ y_i(k) - r_i(k) \end{bmatrix}. \quad (2.4)$$

Z kolei wektor gradientu $\mathbf{g}_k(\mathbf{w}_i)$ jest dany jako:

$$\nabla(g_k(\mathbf{w}_i)) = 2 \frac{\partial g_k(\mathbf{w}_i)}{\partial \mathbf{w}_i^*} = x(k) \frac{[y_i(k) - r_i(k)]^*}{|y_i(k) - r_i(k)|}. \quad (2.5)$$

Zatem, jeśli

$$v_i(k) = y_i(k) - r_i(k), \quad (2.6)$$

to równanie (2.5) można zapisać jako:

$$\nabla(g_k(\mathbf{w}_i)) = \mathbf{x}(k) \frac{v_i^*(k)}{|v_i(k)|}. \quad (2.7)$$

Podstawiając (2.7) w metodzie Gaussa [4], możemy $\mathbf{D}(\mathbf{w}_i)$ zapisać jako:

$$\begin{aligned} \mathbf{D}(\mathbf{w}_i) &= [\nabla(g_1(\mathbf{w}_i)), \nabla(g_2(\mathbf{w}_i)), \dots, \nabla(g_K(\mathbf{w}_i))] \\ &= \left[\mathbf{x}(1) \frac{v_i^*(1)}{|v_i(1)|}, \mathbf{x}(2) \frac{v_i^*(2)}{|v_i(2)|}, \dots, \mathbf{x}(K) \frac{v_i^*(K)}{|v_i(K)|} \right] = \mathbf{X} \mathbf{V}_{iCM}, \end{aligned} \quad (2.8)$$

przy czym:

$$\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)], \quad (2.9)$$

oraz

$$\mathbf{V}_{iCM} = \begin{bmatrix} \frac{v_i^*(1)}{|v_i(1)|} & 0 & \dots & 0 \\ 0 & \frac{v_i^*(2)}{|v_i(2)|} & & \vdots \\ \vdots & & & 0 \\ 0 & \dots & 0 & \frac{v_i^*(K)}{|v_i(K)|} \end{bmatrix}. \quad (2.10)$$

W rezultacie, zestawiając (2.4) i (2.8) otrzymujemy:

$$\mathbf{D}(\mathbf{w}_i)\mathbf{D}^H(\mathbf{w}_i) = \mathbf{X}\mathbf{V}_{iCM}\mathbf{V}_{iCM}^H\mathbf{X}^H = \mathbf{X}\mathbf{X}^H, \quad (2.11)$$

oraz

$$\mathbf{D}(\mathbf{w}_i)\mathbf{g}(\mathbf{w}_i) = \mathbf{X}\mathbf{V}_{iCM} \begin{bmatrix} |v_i(1)| \\ |v_i(2)| \\ \dots \\ |v_i(K)| \end{bmatrix} = \mathbf{X} \begin{bmatrix} v_i^*(1) \\ v_i^*(2) \\ \dots \\ v_i^*(K) \end{bmatrix} = \mathbf{X}\mathbf{v}_i^* \quad (2.12)$$

$$= \mathbf{X}(y_i - r_i)^* \quad (2.13)$$

gdzie

$$\mathbf{v}_i = [v_i(1), v_i(2), \dots, v_i(K)]^T \quad (2.14)$$

$$\mathbf{y}_i = [y_i(1), y_i(2), \dots, y_i(K)]^T \quad (2.15)$$

$$\mathbf{r}_i = [r_i(1), r_i(2), \dots, r_i(K)]^T. \quad (2.16)$$

Wektor y_i jest wektorem danych wyjściowych dla użytkownika i -tego, natomiast r_i jest estymatą formacji elektromagnetycznej informacji przychodzącej od użytkownika i -tego w czasie trwania jednego bitu. Podstawiając równanie (2.11) i (2.13) do równania w metodzie Gaussa otrzymujemy:

$$\begin{aligned} \mathbf{w}_i(l+1) &= \mathbf{w}_i(l) - [\mathbf{X}\mathbf{X}^H]^{-1} \mathbf{X}(y_i(l) - r_i(l))^* = \\ &= \mathbf{w}_i(l) - [\mathbf{X}\mathbf{X}^H]^{-1} \mathbf{X}\mathbf{X}^H \mathbf{w}_i(l) + [\mathbf{X}\mathbf{X}^H]^{-1} \mathbf{X}r_i(l)^* = \\ &= [\mathbf{X}\mathbf{X}^H]^{-1} \mathbf{X}r_i^*(l), \end{aligned} \quad (2.17)$$

gdzie $y_i(l)$ oraz $r_i(l)$ są odpowiednio, wektorem danych wyjściowych oraz estymatą formy elektromagnetycznej sygnału i -tego użytkownika w odniesieniu do wektora wagowego \mathbf{w} w l -tej iteracji. Algorytm LS-DRMTA może adaptować wektory wagowe przy użyciu różnych bloków danych wejściowych w każdej iteracji. Zatem jeśli:

$$\mathbf{X}(l) = [\mathbf{x}(1+lK), \mathbf{x}(2+lK), \dots, \mathbf{x}((l+1)K)] \quad l = 0, 1, \dots, L \quad (2.18)$$

gdzie L jest liczbą iteracji niezbędnych do uzyskania zbieżności algorytmu, zaś K jest liczbą próbek na bit (jeśli wszystkie próbki z okresu trwania jednego bitu są używane do adaptacji), to algorytm LS-DRMTA dla i -tego użytkownika może być opisany przez następujące równania:

$$\mathbf{y}_i(l) = [\mathbf{w}_i^H(l)\mathbf{X}(l)]^T = [y_i(1+lK), y_i(2+lK), \dots, y_i((l+1)K)]^T \quad (2.19)$$

$$\hat{b}_{il} = \text{sgn} \left\{ \text{Re} \left(\sum_{k=1+lK}^{(1+l)K} y_i(k) c_i(k - k_{ri}) \right) \right\} \quad (2.20)$$

$$\mathbf{r}_i(l) = \hat{b}_{il} [c_i(1+lK - k_{ri}), c_i(2+lK - k_{ri}), \dots, c_i((1+l)K - k_{ri})]^T \quad (2.21)$$

$$\mathbf{w}_i(l+1) = [\mathbf{X}(l)\mathbf{X}^H(l)]^{-1} \mathbf{X}(l)\mathbf{r}_i^*(l). \quad (2.22)$$

gdzie $c_i(k)$ jest k -tą próbką rozproszonego sygnału użytkownika i -tego, k_{ri} jest liczbą próbek odpowiednią do opóźnienia sygnału użytkownika i -tego, zaś b_{il} to estymata l -tego bitu dla tego użytkownika.

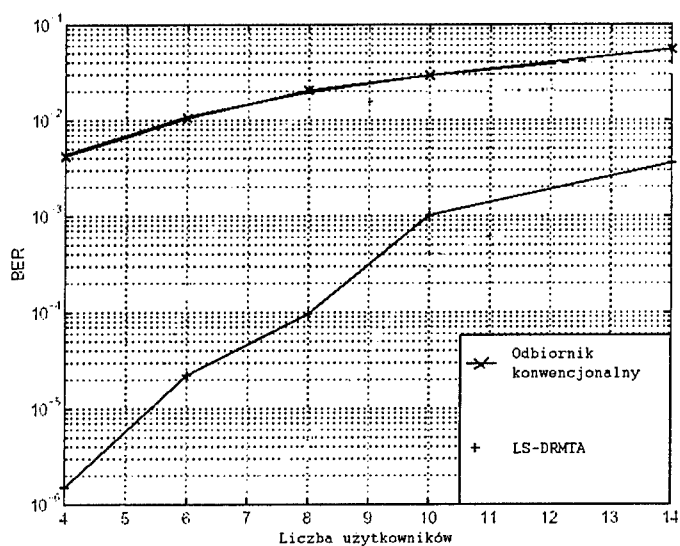
Działanie opisywanego algorytmu można streścić w następujących krokach:

1. Zainicjalizuj wektory wagowe \mathbf{w}_i do \mathbf{w}_p jako p identycznych $M \times 1$ wektorów kolumnowych z pierwszym elementem równym 1 i innymi elementami równymi 0
2. Oblicz wektor wyjściowy szyku używając równania (2.19)
3. Poddaj sygnał użytkownika operacji skupiania i estymuj bit danych według (2.20)
4. Poddaj ponownie rozproszeniu estymowany bit danych za pomocą kodu PN i -tego użytkownika, patrz równanie (2.21)
5. Adoptuj współczynniki wagowe równaniem (2.22)
6. Powtarzaj kroki 2 do 5 dopóki algorytm nie osiągnie konwergencji

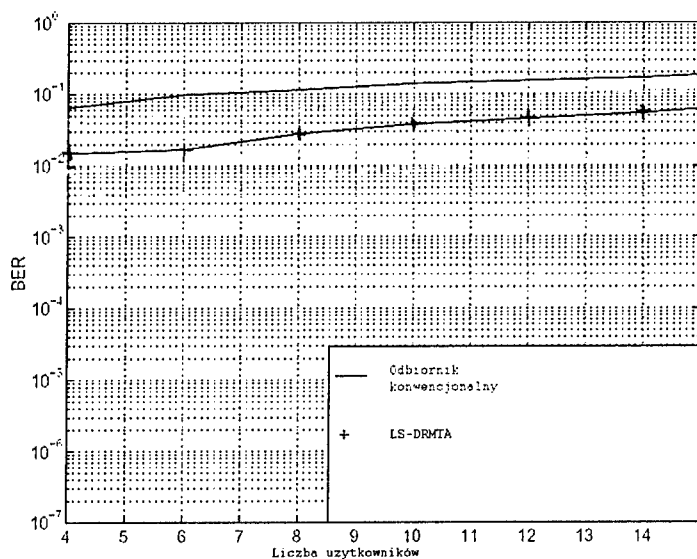
3. WYNIKI SYMULACJI

Do celów symulacji wybrano 2 modele kanału: kanał AWGN oraz kanał wielościeżkowy (wielodrogowy). Badano stopę błędów BER dla zadanej liczby użytkowników. Wybrano wielodostęp DS-CDMA (*Direct Sequence CDMA*) z zyskiem przetwarzania równym 15. Jest to stosunkowo mały zysk przetwarzania, jednakże ma to na celu osiągnięcie pewnych specyficznych celów. Po pierwsze dzięki niewielkiemu zyskowi przetwarzania czas wykonywania symulacji zamknął się w rozsądnych granicach. Po drugie niewielka wartość zysku pozwoliła skoncentrować się na zbadaniu wpływu na system algorytmu adaptacji – i w konsekwencji określenia jaką wartość zysku przetwarzania wniesie stosowanie algorytmu adaptacyjnego, bez zaciemniania obrazu przez znaczne zyski przetwarzania osiągane dla długich sekwencji pseudolosowych. Założono, że system będzie pracował z modulacją BPSK z przepływnością źródłową 128 kbps. Częstotliwość nośną ustalono na 2,05GHz. Zastosowano układ antenowy w postaci matrycy liniowej, złożonej z elementów o tej samej, dookólnej charakterystyce. Liczba elementów układu antenowego wynosiła 8, odległość między elementami równa połowie długości fali. W trakcie trwania 1 chipu próbkowany on był 4 krotnie. Rozmiar bloku danych algorytmu wynosiła 60 próbek, tj. tyle przypadało na jeden nadany bit.

Stosując powyższe założenia otrzymano wykres przedstawiony na rys. 1, obrazujący zależność stopy błędów BER od liczby użytkowników dla kanału AWGN.



Rys. 1. Stopa błędów w funkcji liczby użytkowników dla kanału AWGN, parametry symulacji: $E_b/N_0=8\text{dB}$, kierunki nadejścia sygnału od użytkowników rozmieszczone równomiernie pomiędzy -70° a 90° .



Rys. 2. Stopa błędów w funkcji liczby użytkowników dla kanału wielościeżkowego

Następnie wykonano symulację dla środowiska wielościeżkowego, co oznacza, że w symulacji przyjęto 2 ścieżki propagacji dla sygnału od każdego użytkownika. W wykonanej symulacji przyjęto, że pierwsze ścieżki propagacji są równomiernie rozmieszczone w przedziale kątów nadejścia od -70° do 90° . Kierunek nadejścia sygnału dla drugiej ścieżki propagacji dla każdego użytkownika jest o 10° mniejszy niż ma to miejsce w pierwszej ścieżce. Dodatkowo stosunek mocy sygnału zawartej w pierwszej ścieżce do mocy zawartej w drugiej ścieżce wynosi 0dB, a opóźnienie pomiędzy ścieżkami stanowi 0,5 okresu chipu. Podobnie jak w pierwszej symulacji przyjęto $E_b/N_0=8\text{dB}$. Wyniki przedstawiono na rys. 2.

Wyniki symulacji dla kanału AWGN jak też dla kanału wielościeżkowego pokazują, że rozwijanie algorytmów adaptacyjnych dla sygnałów szerokopasmowych ma uzasadnienie. Dla kanału wielościeżkowego udało się zmniejszyć stopień błędów o cały rząd wielkości, w szerokim zakresie zmienności liczby użytkowników. Potwierdza to konieczność dalszych badań i prób opracowania nowych algorytmów.

4. ZAKOŃCZENIE

W wyniku prowadzonych na świecie prac badawczych w dziedzinie techniki MIMO dotychczasowy dynamiczny rozwój badań nad systemami anten inteligentnych został zahamowany. Dotyczy to w szczególności współpracy adaptacyjnego systemu antenowego z sygnałami szerokopasmowymi, w zastosowaniu do naziemnych systemów radiokomunikacji ruchomej, w tym systemów opartych o wielodostęp DS CDMA i FH CDMA.

Uzyskane wyniki badań stosunkowo prostego algorytmu jakim jest LS-DRTMA wskazują na duży potencjał tkwiący w algorytmach adoptowanych, wywodzących się z klasycznych algorytmów pracy anten inteligentnych. Powstaje pytanie w jakim stopniu algorytmy specjalizowane powstałe wyłącznie z myślą o sygnałach szerokopasmowych są w stanie zwiększyć możliwości stacji bazowej wyposażonej w adaptacyjny szkielet antenowy. Można przypuszczać, że stosowanie odbioru zbiorczego w pewnych środowiskach propagacji spotka się ze znacznymi utrudnieniami i wskazane będzie stworzenie hybrydy układu MIMO i układu adaptacyjnego, w celu dostosowania przyszłych systemów 4G do pracy we wszystkich środowiskach propagacyjnych.

BIBLIOGRAFIA

- [1] Roy R., Kailath T.: *ESPRIT – estimation of signal parameters via rotational invariance techniques*, IEEE Trans. On Acoustics, Speech and Signal Processing, Vol. 37, pp. 984-995, July 1986.
- [2] Hottinen A., Tirkkonen O., Wichman R.: *Multi-antenna Transceiver Techniques for 3G and Beyond*, Wiley 2003
- [3] Widrow B., Stearns S.: *Adaptive Signal Processing* Prentice Hall, 1985.
- [4] Zhigang R.: *Simulation of Adaptive Array Algorithms for CDMA Systems*, Virginia State University, 1996

PERFORMANCE EVALUATION OF LS-DRTMA ALGORITHM IN ANTENNA IMPLEMENTATION

Summary

This article describes the topic of LS-DRMTA adaptation algorithm, which was developed to serve as beamforming method in CMDA systems adaptive antenna arrays. Analitical background was presented, as well as simulation results. On the basis of conducted simulations some final remarks are given.

Wiktor Pawłowski**Katedra Systemów i Sieci Radiokomunikacyjnych, Politechnika Gdańska**

UWARUNKOWANIA PROPAGACYJNE W WSPÓŁCZESNEJ ŁĄCZNOŚCI MORSKIEJ

Streszczenie

W pracy omówiono uwarunkowania propagacyjne z akwenami morskimi objętymi współczesną radiokomunikacją morską. Wskazano na znaczne zróżnicowanie propagacji fal radiowych wynikające z konieczności zapewnienia łączności z różnie odległymi akwenami. Powoduje to wykorzystywanie różnych zjawisk propagacyjnych i stosowanie różnych zakresów fal radiowych w łączności morskiej. Omówiono również uwarunkowania propagacyjne w łączności podwodnej.

1. WPROWADZENIE

Żegluga morska ma zasięg globalny i odbywa się z jednostkami w różnie odległych akwenach. Cechą łączności morskiej jest więc znaczne zróżnicowanie długości tras propagacyjnych. Łączność morska jest realizowana między statkami oraz terminalami na lądzie lub na satelicie. Wyróżnić można łączność przybrzeżną na falach pośrednich i falach ultra-krótkich oraz łączność dalekosiężną krótkofalową, a także łączność satelitarną [1]. Specyficznym przypadkiem łączności morskiej jest porozumiewanie się z odległymi obiektami podwodnymi na falach bardzo długich.

Ze względu na zależność właściwości propagacyjnych fal radiowych od częstotliwości oraz długości trasy jest konieczne w łączności morskiej wykorzystanie praktycznie wszystkich zakresów fal radiowych, od fal bardzo długich po mikrofałe.

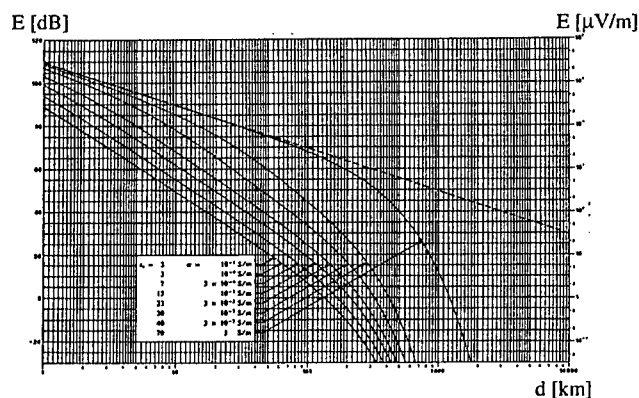
2. ŁĄCZNOŚĆ PRZYBRZEŻNA

W łączności morskiej wykorzystywane są oba rodzaje fali przyziemnej, fala powierzchniowa, która powstaje gdy antena (w postaci na przykład masztu promieniującego) ustawiona jest bezpośrednio na powierzchni Ziemi oraz fala przestrzenna powstająca, gdy antena jest umieszczona przynajmniej na wysokości 10λ nad powierzchnią Ziemi.

2.1. Propagacja fali powierzchniowej [2]

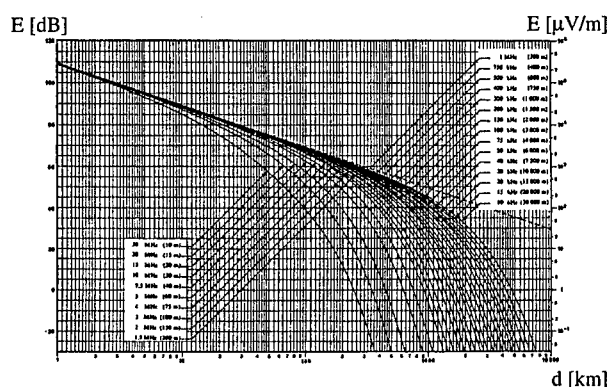
W łączności przybrzeżnej wykorzystywane jest pasmo 2 MHz z zakresu fal pośrednich obejmujących częstotliwości od 1.5 do 4 MHz. Praktyczne znaczenie ma fala powierzchniowa o polaryzacji pionowej. Wartość natężenia pola fali powierzchniowej zależy do strat w Ziemi powstałych na drodze propagacji, które na trasie morskiej są mniejsze niż na lądzie.

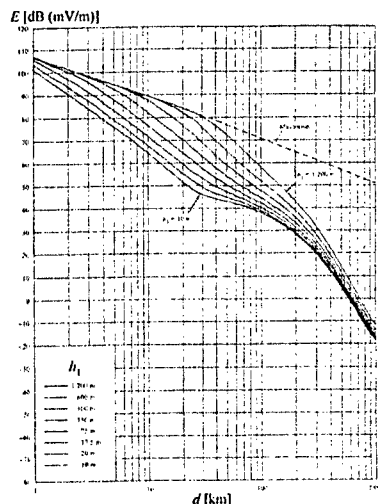
Na ogół wartość natężenia pola fali powierzchniowej wyznaczana jest na podstawie uzyskanych na drodze teoretycznej krzywych propagacji fali powierzchniowej (spolaryzowanej pionowo) pokazujących przebieg wartości natężenia pola w funkcji odległości. Założono przy tym, że antena krótka zasilana jest mocą 1 kW. Parametrem są stała dielektryczna i przewodność właściwa podłoża trasy. Krzywe opracowane zostały dla wybranych częstotliwości z zakresu 10 kHz do 30 MHz. Przykłady krzywych, dla częstotliwości 2.2 MHz z uwzględnienia wpływu parametrów podłoża pokazano na rys.1, zaś na rys.2. krzywe dla trasy morskiej – z uwzględnieniem wpływu częstotliwości.



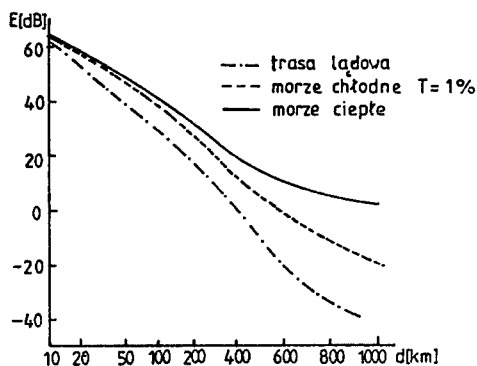
Rys.1. Krzywe propagacji przy częstotliwości 2,2MHz

Za pomocą tego sposobu propagacji realizowana jest łączność w akwenach rozciągających się od około 30 MM do około 150 MM od brzegu. Jest to tak zwany akwen A2. Omawiany zakres częstotliwości charakteryzuje się dużymi poziomami szumów różnego pochodzenia. Szczególne znaczenie mają szumy pochodzące od wyładowań atmosferycznych. Zakłócenia te powstają w czasie burz w strefie tropikalnej i rozchodzą się za pośrednictwem refrakcji jonosferycznej na znaczne odległości, w naszej strefie klimatycznej objawiając się w postaci szumów, tzw. atmosferycznych. Wpływają one w istotny sposób na jakość łączności morskiej. Z tego względu prognozowanie poziomu szumów stanowi ważny element metody przewidywania zasięgów przy założonej jakości łączności.





Rys. 3. Przykładowe statystyczne krzywe propagacji, $f = 100$ MHz, $t = 0,01$; morze chłodne



Rys. 4. Porównanie krzywych dla $T=1\%$ dla trasy lądowej i tras morskich

Odrębny przypadek stanowi propagacja na trasie mieszanej lądowo-morskiej. Ze względu na odmienne warunki propagacji nad lądem i morzem trasę taką powinno się traktować jako niejednorodną [4]. W łączności morskiej anteny nabrzeżne umieszczone są blisko brzegu morskiego przypadek ten nie odgrywa zasadniczej roli.

Za pomocą omówionego sposobu propagacji fal radiowych realizowana jest łączność morska w pasie sięgającym do około 30 MM od brzegu. Jest to tak zwany akwen A1. Wykorzystywane jest pasmo 156 MHz. Długości tras propagacji wynikają z quasi-horyzontowego zasięgu zależnego od wysokości umieszczenia anten, przede wszystkim brzegowych. Omówiony sposób propagacji wykorzystywane jest także przy bezpośredniej łączności statkami.

3. ŁĄCZNOŚĆ ZA POŚREDNICTWEM SATELITÓW

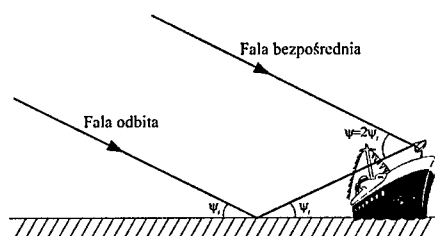
Morskie łącze satelitarne jest szczególnym przypadkiem łącza geostacjonarnego. O jego specyfice decydują warunki w jakich znajduje się terminal morski – statek realizując rejs porusza się na pofalowanej, dobrze odbijającej fale radiowe powierzchni morza. Ze względu na stosowanie na statkach stosunkowo mało kierunkowych anten istnieją warunki do wystąpienia propagacji wielodrogowej [5]. Satelitarna łączność morska realizowana jest za pośrednictwem systemu INMARSAT. Obejmuje ona tzw. akwen A3, czyli wszystkie akweny morskie znajdujące się w pasie między równoleżnikami 70N i 70S.

3.1. Tłumienie propagacyjne

Satelity geostacjonarne umieszczone są na orbicie o wysokości 35.810 km. Z długich tras propagacji wynikają duże wartości strat propagacyjnych. Są to tzw. podstawowe straty transmisji, które wystąpiłyby gdyby propagacja odbywała się w wolnej przestrzeni. Biorąc pod uwagę, że komunikacja odbywa się w paśmie L, podstawowe straty transmisji wy-

noszą w przypadku łącza satelita – stacja okrętowa oraz stacja okrętowa – satelita odpowiednio 188.5 dB oraz 189 dB. Inne efekty powodujące straty transmisji takie, jak np. tłumienie w hydrometeorach, są przy częstotliwości 1.5 GHz jeszcze słabe. Na peryferiach systemu, przy kątach elewacji 5 i 10 stopni, szacuje się je odpowiednio na 0.4 dB i 0.2 dB. Efekty wywołane przez zjawiska w jonosferze są w paśmie L do zaniedbania.

3.2. Propagacja wielodrogowa



Rys.5. Warunki pracy terminalu okrętowego

okrętowej. Zakłada się, że fala odbita od morza odbierana jest główną wiązką anteny, której charakterystykę można opisać następującym wyrażeniem.

$$G = -4 \cdot 10^{-4} (10^{G_m/10} - 1) \Psi^2, \quad (1)$$

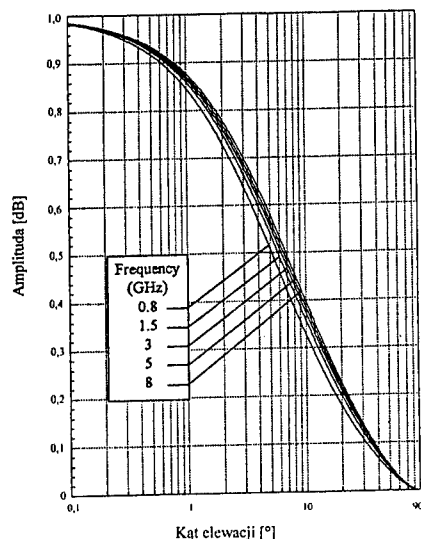
gdzie Ψ jest kątem elewacji głównej wiązki anteny, zaś G_m maksymalną wartością zysku anteny.

Falę odbitą od powierzchni morza określa się korzystając z wzoru Fresnela. Na rys.6. przytoczono przebieg modułu współczynnika odbicia w zależności od kąta elewacji w stopniach.

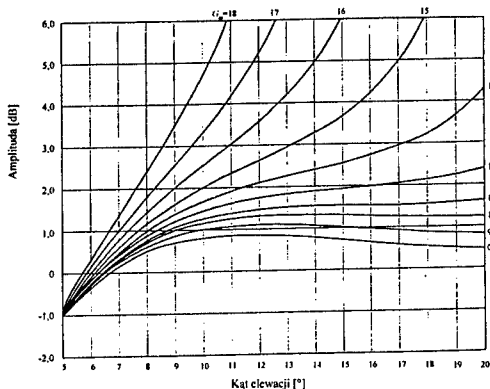
Średnią moc fal odbitych od powierzchni morza, P_r , odniesioną do mocy promieniowania bezpośredniego można wyrazić następująco

$$P_r = G + R + \eta_l \quad (2)$$

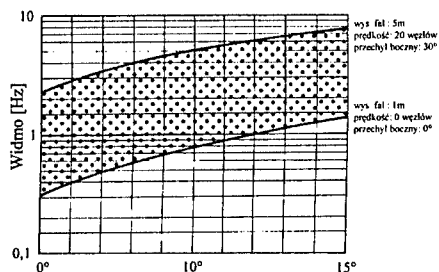
gdzie: G [dB] – zysk na kierunku nadejścia fali, $R = 20 \log |R_l|$, przy czym $|R_l|$ moduł współczynnika odbicia, zaś η_l jest współczynnikiem podającym stosunek składowej rozproszonej do odbitej od powierzchni morza. Jest to tak zwany uśredniony znormalizowany współczynnik rozproszenia – patrz rys.7. Chcąc wyznaczyć głębokość zaników należy skorzystać z rozkładu Nakagami-Rice'a.



Rys.6. Przebieg modułu współczynnika odbicia



Rys.7. Przebieg uśrednionego znormalizowanego współczynnika rozproszenia



Rys.8. Zakres zmian widma zaników w funkcji kąta elewacji głównej wiązki anteny okrętowej

Widmo zaników spowodowanych wielodrogowością propagacji rośnie ze wzrostem wysokością fal morskich, szybkością statku oraz natężeniem jego kotłusań. Słabą zależność stwierdza się od polaryzacji fali, a także od zysku anteny (przy zysku mniejszym od 10 dB).

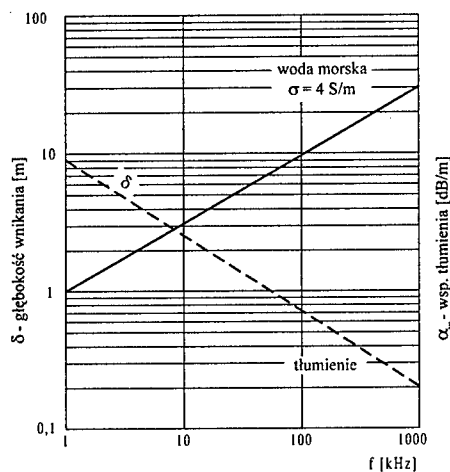
Na rys.8 pokazano w funkcji kąta elewacji głównej wiązki anteny okrętowej przewidywany zakres zmian widma zaników – obszar zakropkowany na rysunku. Granice tego obszaru określone są typowymi dla żeglugi morskiej zakresami zmian wysokości fal morskich od 1 do 5 metrów, prędkości statku od 0 do 20 węzłów oraz przechyłów bocznych od 0 do 30 stopni.

4. SPOSOBY PROPAGACJI PRZY ŁĄCZNOŚCI Z OBIEKTAMI PODWODNYMI [6], [7]

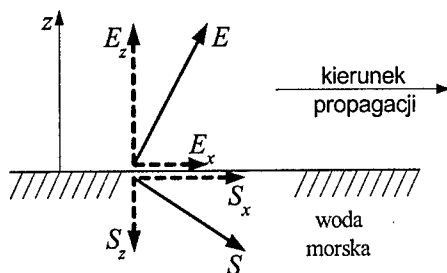
Fale radiowe stosowane do łączności z obiektami podwodnymi powinny rozchodzić się stosunkowo dobrze w wodzie. Z drugiej strony trzeba wziąć pod uwagę duże odległości, rzędu tysięcy kilometrów, w których położone mogą być obiekty podwodne. Warunki te spełniają fale zakresu ELF, tzw. myriametrowe. Głębokości wnikania w wodzie morskiej pokazano na Rys.9 (woda morska jako dobry przewodnik jest środowiskiem, w które fale radiowe słabo wnikają).

W praktyce radiokomunikacyjnej wykorzystuje się na ogół fale spolaryzowane pionowo. W związku z tym, że woda morska jest środowiskiem stratnym, wektor natężenia pola elektrycznego nad morzem jest pochylony, tzn. pojawia się składowa pozioma, jak widać to na rys.10. Składowa ta, reprezentująca przy komunikacji nawodnej straty, wykorzystywana jest do komunikowania się z obiektem podwodnym. Nad wodą w miejscu przebywania obiektu podwodnego należy wytworzyć pole elektryczne o natężeniu zapewniającym łączność między powierzchnią morza a obiektem podwodnym. Do wytworzenia dostatecznie dużego natężenia pola są wykorzystywane dwa sposoby propagacji fal myriametrowych: fala powierzchniowa oraz propagacja w falowodzie jonosferycznym.

Propagacja fali powierzchniowej fal bardzo długich zasadniczo nie różni się od propagacji tej fali w przypadku fal pośrednich, omówionej już w punkcie 2.1.



Rys.9. Głębokość wnikania fali



Rys.10. Układ wektorów fali nad morzem

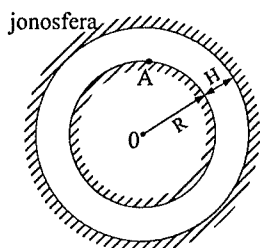
Przy dużych długościach tras propagacji, większych od 1500 km, wartości natężenia pola uzyskane na drodze pomiarowej są większe od wynikających z krzywych propagacji. Zjawisko to wyjaśniono stawiając tezę, że fale myriametrowe rozchodzą się w falowodzie kulistym, który powstaje w obszarze między powierzchnią Ziemi a dolną warstwą jonosfery, pokazanym na rys.11. Uproszczona analiza jonosferycznego mechanizmu falowodowego zakłada, że powierzchnie ograniczające dobrze odbijają fale radiowe. W takich warunkach natężenie pola zmniejsza się jedynie na skutek malenia gęstości powierzchniowej mocy, podobnie jak ma to miejsce w wolnej przestrzeni. Mechanizm ten objaśniono na rys.12. Zakłada się, że antena o charakterystyce dookólnej znajduje się w punkcie A. Powierzchnia, S , na którą rozkłada się energia niesiona przez fale, jest częścią powierzchni bocznej stożka o osi OB i tworzącej OC , ograniczonej kulami o promieniach R oraz $R + H$. Biorąc pod uwagę, że $H \ll R$ uzyskuje się wzór przybliżony

$$S = 2\pi \left(R + \frac{H}{2} \right) H \sin \Theta, \quad (3)$$

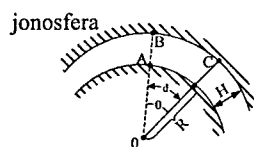
gdzie: R – promień Ziemi, H – wysokość falowodu oraz θ – kąt środkowy na wielkim kole kuli ziemskiej od powiadający długości trasy d .

Przyjmując, że moc promieniowania wynosi P_r , otrzymuje się, przy powyższych założeniach, następujące wyrażenie na natężenie pola w falowodzie jonosferycznym

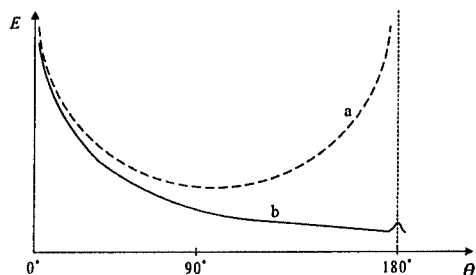
$$E = \sqrt{\frac{60P_r}{\left(R + \frac{H}{2}\right)H \sin \theta}} \quad (4)$$



Rys.11. Falowód kulisty



Rys.12. Objasnienie propagacji falowodowej



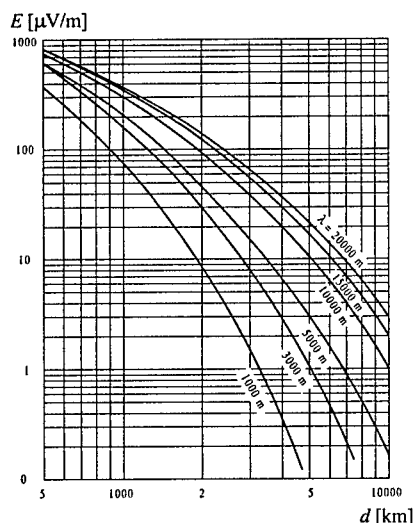
Rys.13. Natężenie pola w falowodzie

Wartości natężenia pola w przedziale $0 < \theta < 90^\circ$ są malejące. Potem rosną, jak pokazuje to linia przerywana na rys.13. Jest to wynikiem tzw. efektu antypodowego, polegającego na lokalnym maksimum natężenia pola w rejonie antypodu – linia ciągła na rys.13.

Porównując oba przytoczone mechanizmy propagacji fal myriametrowych można stwierdzić, że przy tzw. „małych odległościach”, tj. do około 1000 km przeważa fala powierzchniowa – warunki propagacji można przewidywać w oparciu o teoretyczne krzywe propagacji fali powierzchniowej. Przy większych odległościach dominuje fala jonosferyczna. Ścisłe metody analizy warunków propagacji są w tym przypadku skomplikowane. Często korzysta się w tym przypadku z empirycznego wzoru Austin'a-Cohen'a (patrz rys. 14)

$$E = \frac{300\sqrt{P_r}}{d} \sqrt{\frac{\theta}{\sin \theta}} e^{\frac{0,0014d}{\lambda^{0.6}}} \quad [\text{mV/m}], \quad (5)$$

gdzie: P_r – moc promieniowana [kW], d – długość trasy [km] oraz λ – długość fali [km].



Rys.14. Krzywe propagacji wg Austina-Cohena

Ze względu na obszerność tematyki dotyczącej uwarunkowań propagacyjnych w łączności morskiej i ograniczone ramy pracy, skoncentrowano się w niniejszym opracowaniu przede wszystkim na najistotniejszych zjawiskach propagacyjnych mających znaczenie w eksploatowanych obecnie systemach. Pominieto zagadnienia dotyczące propagacji krótkofalowej. Zakłada się, że propagacja krótkofalowa powinna być wykorzystywana jedynie przy łączności z obiektami pływającymi na akwenach morskich położonych poza pasmem między równoleżnikami 70N – 70S, czyli na obszarach nie objętych radiokomunikacją satelitarną. Jest to tzw. akwen A4 [8].

BIBLIOGRAFIA

- [1] Katulski R.J., Kosiński A., Pawłowski W., Żurek J.: *Stan obecny i perspektywy radiokomunikacji morskiej*, Przegląd Telekomunikacyjny, 12/1999, 816-820.
- [2] Pawłowski W.: *Uwarunkowania szumowe w pośredniofalowej radiokomunikacji morskiej*, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, Wrocław, 25 – 27 czerwca 2003, 433-436.
- [3] Pawłowski W.: *Uwarunkowania propagacyjne łączności z jednostkami morskimi w przyszłościowych systemach radiokomunikacyjnych*, Krajowa Konferencja Radiotelekomunikacji, Radiofonii i Telewizji, Poznań, 14 – 16 maja 2001, 6.3-1 do 6.3-4
- [4] Pawłowski W.: *Modele propagacyjne dla radiokomunikacji morskiej w świetle dokumentów ITU-R*, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, Gdańsk, 12 – 14 czerwca 2002, 101 - 104
- [5] Pawłowski W., *Analiza łączności satelitarnej radiokomunikacji morskiej*, V Krajowa Konferencja Radiodyfuzji i Radiokomunikacji KKRR'2000, Poznań 6 – 8 czerwca 2000, 307 - 310
- [6] Katulski R.J., Kosiński A., Pawłowski W., Studański R.: *Podstawy teorii i techniki łączności podwodnej*, Przegląd Telekomunikacyjny, 9/1998, 621-629.
- [7] Pawłowski W.: *Mechanizmy propagacji fal bardzo długich w atmosferze ziemskiej stosowanych w wybranych systemach radiokomunikacji morskiej*, Krajowa Konferencja Radiodyfuzji i Radiokomunikacji KKRR'98, Poznań, 18 – 20 maja 1998, 287 – 290
- [8] Pawłowski W., *Uwarunkowania propagacyjne w radiokomunikacji morskiej*, Krajowa Konferencja Radiokomunikacji Ruchomej KKRR'97, Poznań, 26 – 28 maja 1997, 210 - 213

PROPAGATION CONDITIONS IN PRESENT MARITIME COMMUNICATION**Summary**

In the paper the most important phenomena affecting the propagation conditions of radio waves in maritime communication are described. Propagation conditions of the ground wave and sky wave, as well as in satellite systems, are discussed. The problems connected with the propagation with underwater objects, are presented, too.

Marek Przyborski

Zakład Łączności Akademii Marynarki Wojennej

KONCEPCJA SYSTEMU TRANSMISJI DANYCH DLA POTRZEB AKCJI RATOWNICZYCH

Streszczenie

W artykule zaprezentowano nowy kierunek badań w zakresie przekazywania informacji w czasie akcji ratowniczej. Uzasadniono konieczność wprowadzania takiego systemu, przedstawiono założenia jakim musi odpowiadać oraz schemat obiegu informacji w trakcie akcji ratowniczej i graf dojścia szybkiej łodzi ratowniczej do rejonu poszukiwań.

1. WSTĘP

Wykorzystanie akwenów morskich do żeglugi, od wieków umożliwiało człowiekowi wymianę gospodarczą, nawiązywanie kontaktów i poznawanie świata. Dziś szeroko rozumiana gospodarka morska, traktowana w wielu państwach jako gałąź przemysłu, angażuje wielu ludzi. Dlatego też zapewnienie należytego bezpieczeństwa na wodach mórz i oceanów, należy traktować jako sprawę priorytetową.

Głównym zadaniem działań ratowniczych SAR jest zaplanowanie poszukiwań, a następnie skuteczne zlokalizowanie i wyratowanie, obiektu który znalazł się w sytuacji zagrożenia na morzu. Z tego też powodu od wielu lat poszukuje się najbardziej wydajnego sposobu na precyzyjne lokalizowanie zagrożonego obiektu. Wizualizacja działań w czasie rzeczywistym oraz szacowanie skuteczności wykonanego zadania byłoby wspaniałym narzędziem umożliwiającym zwiększenie efektywności akcji ratowniczych.

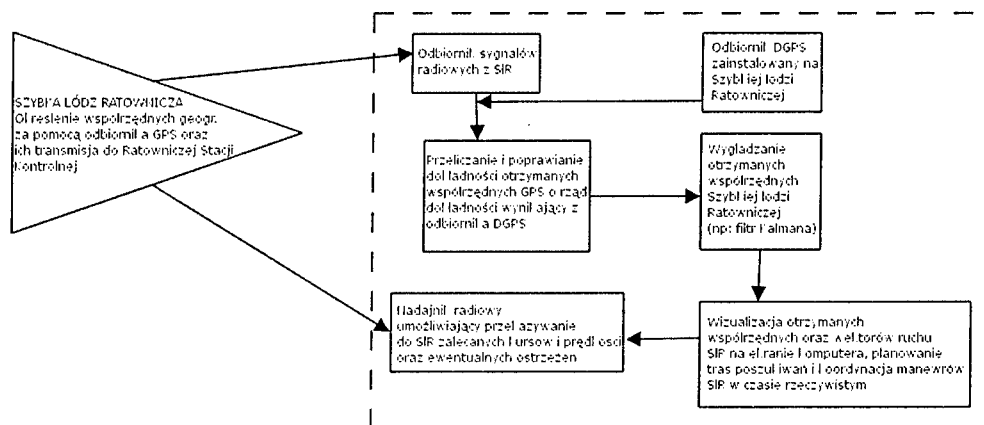
Jednym z niezbędnych elementów używanym w akcjach ratowniczych na morzu są szybkie łodzie ratownicze [1,3,4], które mogą stanowić wyposażenie większych jednostek ratowniczych lub też być wodowane z łądu jeżeli miejsce katastrofy znajduje się blisko brzegu. Aby zapewnić optymalne wykorzystanie tych oraz innych środków ratowniczych należy zapewnić monitorowanie pozycji oraz transmitowanie tych danych do ośrodka koordynacyjnego na brzegu. Informacja o pozycji łodzi czy też okrętu ratowniczego usprawniłaby znacząco kierowanie akcją ratowniczą. Osoba odpowiedzialna za kierowanie akcją z łądu miałaby pewną informację i nie musiałaby opierać się na meldunkach radiowych pochodzących od załogi łodzi.

Wizualizacja tych danych mogłaby zaowocować dokładniejszym przeszukaniem tzw. obszaru najbardziej prawdopodobnego, a w rezultacie także wykonanie zadania w relatywnie krótszym czasie.

Wszystkie powyższe czynniki są więc bardzo istotne w procesie podejmowania akcji ratowania zagrożonego życia na morzu, co w konsekwencji może prowadzić do oczekiwanej skuteczności takich działań. System transmisji danych nawigacyjnych na brzeg będzie jednym z elementów szerszego systemu, którego celem będzie stworzenie struktur do opracowania nowych bardziej efektywnych technik poszukiwania i ratowania życia ludzkiego na morzu.

2. ZAŁOŻENIA NA SYSTEM TRANSMISJI DANYCH

Głównym celem jest opracowanie systemu, koordynowania działań łodzi po osiągnięciu rejonu katastrofy. System transmisji danych o pozycji jest tylko jednym z elementów większej całości, na którą będą się składały moduły wizualizacji sytuacji w rejonie wypadku, wspomaganie podejmowania decyzji, obliczania optymalnych scenariuszy przeszukiwania rejonów. Wszystkie te przedsięwzięcia zmierzają do tego, by w możliwie krótkim czasie odnaleźć rozbitków. Koncepcja obiegu informacji w systemie została przedstawiona poniżej na rys.1.

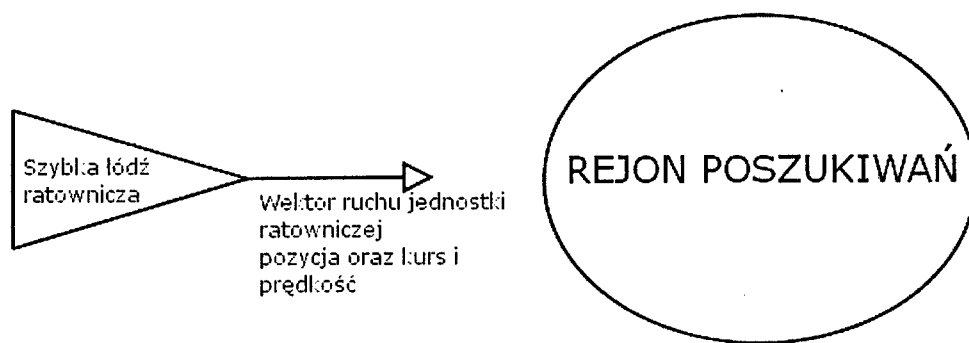


Rys.1. Schemat obiegu informacji w proponowanym systemie koordynowania działań łodzi ratowniczych w trakcie akcji ratunkowej na morzu.

Środkiem wiodącym do realizacji powyżej przedstawionego celu będzie wykonanie serii badań empirycznych, mających na celu przeanalizowanie możliwości zastosowania odbiorników pozycyjnych satelitarnych systemów nawigacyjnych zamontowanych na szybkich łodziach ratowniczych oraz wykorzystanie różnych systemów radiowych do transmisji danych o pozycji jednostki. Podczas tworzenia systemu należy uwzględnić takie aspekty jak:

- koszt transmisji,
- szybkości transmisji,
- dostępności systemu,
- niezbędny zasięg łączności,
- gabaryty aparatury

Jednym z ważnych problemów które należy rozważyć jest określenie możliwości zastosowania segmentu komunikacyjnego, który jest elementem składowym coraz popularniejszego i szukającego wciąż nowych zastosowań europejskiego systemu satelitarnego GALILEO. Wśród rozważanych systemów radiowych znajdują się sieci telefonii komórkowej a w nich usługi takie jak GPRS i SMS. W systemach łączności radiowej powinny być zastosowane odpowiednie radiomodemy. Koszty transmisji jak również samych urządzeń są dość ważnym kryterium. Jeżeli byłaby szansa wynegocjowania porozumienia np. z operatorem telefonii komórkowej, które umożliwiłoby zredukowanie kosztów transmisji GPRS lub usługi SMS, wówczas mogłoby to stanowić ofertę konkurencyjną w stosunku do innych urządzeń radiowych. Bardzo ważnym w przypadku telefonii komórkowej jest wielkość telefonów, możliwość zintegrowania ich z komputerem oraz prostota obsługi. Rozważane są dwa warianty konfiguracji urządzeń do określania pozycji i jej transmisji do Centrum Koordynacyjnego na lądzie. Zestaw na stałe zamontowany na łodzi ratowniczej oraz drugi przypadek gdy zestaw będzie zabierany przez załogę do łodzi tylko na okres trwania akcji. Wielkość takiego zestawu urządzeń i możliwość jego miniaturyzacji byłaby jednym z bardzo ważnych kryteriów.



Rys. 2. Schemat dotarcia łodzi ratowniczej do rejonu poszukiwań.

Uwzględnienie parametrów ruchu jednostki ratowniczej [2,5], jest również ważnym elementem. Ma to ogromny wpływ na interwał z jakim należałoby uaktualniać pozycję (rys. 2) ponieważ ciągła transmisja stanowiłaby znaczne obciążenie dla urządzeń radiowych. Wybór interwału jest bardzo trudny ponieważ szybkości z jakimi może poruszać się łódź ratownicza są dość znaczne (rzędu nawet 60 do 70 węzłów) więc wydłużony czas uaktualniania pozycji skutkowałby znacznymi przekłamaniami w pozycji. Natomiast podczas poruszania się wśród rozbitków unoszących się na wodzie prędkość będzie bardzo mała więc i czas uaktualniania powinien być dostosowany do tych warunków (patrz rys. 2).

3. PODSUMOWANIE

Ratowanie życia ludzkiego na morzu jest nieodłącznym elementem związanym z przemieszczaniem się statków po morzach oraz eksploracją zasobów jakie znajdują się w oceanach. Stąd też działania te mają bardzo wysoki priorytet. Sprawność i efektywność działań ratowniczych w dużym stopniu zależy od czasu w jakim siły ratownicze mogą znaleźć się na miejscu zdarzenia oraz od koordynacji całej akcji. Szybkie łodzie ratownicze, wodowane z większych jednostek czy też bezpośrednio z brzegu, dzięki swoim zdolnościom manewrowym umożliwiają szybkie dotarcie na miejsce katastrofy, a wyposażenie w środki umożliwiające sprawne kontrolowanie ich działań z brzegu sprawiłoby iż stałyby się bardzo sprawnym środkiem podczas kierowania akcjami ratowniczymi. Ponieważ nasze siły ratownicze przede wszystkim przeznaczone są do udzielania pomocy w naszej strefie ekonomicznej więc istnieje możliwość wykorzystania różnych systemów nawigacyjnych, które są dostępne na wodach Bałtyku Południowego. Przeprowadzona w początkowej fazie projektu analiza teoretyczna i praktyczna, umożliwi określenie najbardziej wydajnego systemu pozycjonowania, który następnie przyjęty zostanie do kolejnej fazy, gdzie to realizowana będzie współpraca pomiędzy systemem nawigacyjnym, a systemem radiowym transmitującym dane nawigacyjne do Centrum Koordynacyjnego na lądzie.

Kolejną fazą realizacji projektu, będzie wypracowanie optymalnego rodzaju taktyki poszukiwania zagrożonych obiektów na morzu. Zakłada się, że zostanie zastosowany satelitarny system pozycjonowania, oraz system transmisji danych nawigacyjnych.

BIBLIOGRAFIA

- [1] ATP-10(D) *Poszukiwanie i ratownictwo. Instrukcja sprzymierzonych*,
- [2] Bednarczyk M., Pyrchla J., Stateczny A.: *Lokalizacja wypadku morskiego w systemie SAR – próba formalizacji problemu*. Zeszyty naukowe, WSM Szczecin 2000.
- [3] Drogosiewicz M. Haffner J.: *Ratownictwo morskie. Elementy ratownictwa na morzu*, WSMW, Gdynia 1986.
- [4] Duda D. : *Ratowanie życia ludzkiego na morzu*, WSM, Gdynia 1990.
- [5] Morse M.P., Kimball E.G.: *Methods of operations research*. Military Operations Research Society, Alexandria, Virginia 1998.

CONCEPTION OF THE DATA TRANSMISSION SYSTEM FOR THE NEEDS OF SEARCH AND RESCUE ACTIONS

Summary

In the paper the new direction of the research in the field of data transmission during search and rescue action has been presented. The explanation of the introduction such a system as well as the scheme of the data flow during action and the graph of approaching of the boat to the search area.

Ryszard Studański

Instytut Podstaw Techniki, Akademia Marynarki Wojennej, Gdynia

TRENDY ROZWOJOWE TRANSMISJI DANYCH WE WSPÓŁCZESNEJ MORSKIEJ RADIOKOMUNIKACJI WOJSKOWEJ

Streszczenie

W artykule przedstawiono trendy rozwojowe transmisji danych we współczesnej morskiej radiokomunikacji wojskowej. Opisano istotę działania systemu radiokomunikacyjnego stosującego technikę szybkiego hoppingu częstotliwościowego z niedeterministycznym wybieraniem nośnych. Zaproponowano sposób przeciwdziałania szumowi pasmowemu oraz zakłóceniom wieloczęstotliwościowym w celu uzyskania możliwie wysokiej jakości pracy systemu.

1. WSTĘP

Działanie pojedynczych okrętów jak też zespołów okrętów nawodnych na morzu narażone jest, w przypadku ich rozpoznania, na skuteczne oddziaływanie przeciwnika. Brak naturalnych osłon na otwartym morzu powoduje, że zapewnienie skrytości działania jest jednym z kluczowych warunków uzyskania sukcesu militarnego. Systemy radiokomunikacyjne są niestety czynnikiem ujawniającym pozycje okrętów, a jednocześnie są one praktycznie jedynym sposobem zapewnienia łączności okrętom ze sztabami wyższych szczebli dowodzenia. Dlatego też trwają prace nad systemami łączności, od których oczekuje się wysokiej odporności na zakłócenia celowe oraz takich rozwiązań, by lokalizacja źródeł emisji sygnałów była utrudniona.

2. CHRAKTERYSTYKA WSPÓŁCZESNYCH SYSTEMÓW RADIOKOMUNIKACJI MORSKIEJ

Obecnie na morzu są stosowane techniki łączności wykorzystujące wąskopasmowe kanały radiokomunikacyjne i modulacje cyfrowe. Rozpowszechnione jest automatyczne zestawianie połączeń z wykorzystaniem urządzeń DSC (ang. *Digital Selective Connection*). Do zabezpieczenia przed błędami wprowadzanymi przez kanał jest stosowane kodowanie korekcyjne FEC (ang. *Forward Error Correction*) wspomagane kodowaniem detekcyjnym i techniką ARQ (ang. *Automatic Repeat Request*), jeśli w grę wchodzi transmisja danych. W zakresie fal krótkich wybór częstotliwości nośnej z przyznanego

zbioru, jest dla danego połączenia radiowego wyznaczany automatycznie za pomocą przystawki automatycznego zestawiania linii radiowej ALE (ang. *Automatic Link Establishment*) na podstawie danych o aktywności słonecznej, porze roku, porze doby i lokalizacji wzajemnej użytkowników. W celu przeciwdziałania zanikom i zakłóceniom addytywnym najczęściej stosowany jest powolny hopping częstotliwościowy. W wojskowych systemach radiokomunikacyjnych jakość transmisji danych jest oceniana na bieżąco, a w przypadku jej pogorszenia poniżej zakładanego progu stopy błędów automatycznie następuje zmiana organizacji transmisji danych lub zmiana częstotliwości nośnej.

W ostatnich latach obserwuje się wzrost zainteresowania zastosowaniem w systemach radiokomunikacyjnych nowych technik rozpraszania widma sygnałów. Wynika to z dużej odporności tych technik na zaniki sygnałów i interferencje. Technika zwielokrotnienia kodowego w użytkowaniu wspólnego kanału radiowego, związana z bezpośrednim rozpraszaniem widma sygnałów DS CDMA (ang. *Direct Sequence Code Division Multiple Access*), z samej swojej natury dostarcza rozwiązań, które odznaczają się wyróżniającą się odpornością na zakłócenia w porównaniu z innymi technikami wspólnego użytkowania kanału radiowego. Niemniej w warunkach pracy morskich, wojskowych systemów radiokomunikacyjnych, realizujących połączenia rozproszone, gdy nie wchodzi w grę łączność z wykorzystaniem stacji bazowej i rozgraniczenie ruchu „w górę” i „w dół”, zastosowanie tej techniki wydaje się być ograniczone. Natomiast technika hoppingu częstotliwościowego FH (ang. *Frequency Hopping*) jest w praktyce stosowana w wersji powolnego hoppingu (SFH ang. *Slow Frequency Hopping*) zarówno w zastosowaniach militarnych (np. w radiostacji PR4G) jak też w systemach telefonii komórkowej GSM (ang. *Global Mobile System*) i jej zalety są znane lecz współcześnie nie są już wystarczająco korzystne. Od początku lat 80-tych zaczęły pojawiać się pierwsze publikacje [1], [2] na temat tzw. szybkiego hoppingu częstotliwościowego (ang. *Fast Frequency Hopping* – FFH), który polega na przesyłaniu sygnałów elementarnych kolejno na wielu częstotliwościach nośnych. W literaturze opisywane są systemy FFH:

- z deterministycznym wybieraniem częstotliwości nośnych, w których odbiornik synchronicznie z nadajnikiem zmieniają częstotliwości nośne;
- z niedeterministycznym wybieraniem częstotliwości nośnych, w których wybór nośnej zależy od postaci przesyłanego sygnału elementarnego (symbolu) i adresu odbiornika.

W referacie przedstawiono ocenę jakości systemu z niedeterministycznym wybieraniem częstotliwości nośnych, który wydaje się szczególnie atrakcyjny dla zastosowań wojskowych. Szczegółowy opis pracy tego systemu przedstawiono w publikacjach [3], [4] i [5].

3. CHARAKTERYSTYKA SYSTEMU FFH CDMA

System FFH z niedeterministycznym wybieraniem nośnych wykorzystuje pewną liczbę M częstotliwości nośnych, która zarazem odpowiada liczbie możliwych symboli. Ważne jest, aby odstęp częstotliwości pomiędzy nośnymi był większy niż pasmo koherencji, wówczas zaniki występujące na poszczególnych nośnych można traktować jako niezależne procesy stochastyczne. Każdy odbiornik posiada indywidualny adres, który jest wektorem L wymiarowym, a składowe wektora są liczbami całkowitymi ze zbioru $\{0, \dots, M-1\}$. Symbol jest przesyłany kolejno na L różnych częstotliwościach nośnych. Wybór l -tej częstotliwości nośnej zależy od postaci symbolu odpowiadającej określonej wartości ze zbioru $\{0, \dots, M-1\}$ oraz l -tej składowej wektora adresu odbiornika, gdyż tak określona wartość symbolu jest sumowana (według reguły modulo M) z każdą składową wektora adresu. W konsekwencji, powstaje wektor sygnałów nadanych, również L wymiarowy, którego składowe wyznaczają kolejno numery częstotliwości z banku częstotliwości noś-

ných. Nadajniki innych współużytkowników, korzystając oczywiście z tego samego zbioru częstotliwości nośnych, mogą emitować sygnały jednocześnie na różnych lub tych samych częstotliwościach nośnych. Wycinek nośnej reprezentujący każdą składową wektora sygnałów nadanych nazywany jest chipem. Odbiornik przez czas trwania symbolu dokonuje detekcji sygnałów jednocześnie na wszystkich częstotliwościach nośnych. W tym celu w odbiorniku jest tworzona tablica zerojedynkową $M \times L$ (jedynka logiczna w tablicy oznacza detekcję sygnału na określonej nośnej). Przypisując każdemu niezerowemu elementowi wartość odpowiadającą numerowi wiersza pomniejszoną o jeden, można dokonać odejmowania, każdej niezerowej składowej tablicy odebranych sygnałów z współrzędną wektora adresu o tych samych indeksach kolumnowych. W ten sposób powstaje tablica odtworzonych sygnałów nadanych, w której takie same wartości składowych w każdej z kolumn tablicy wskazują na postać odebranego symbolu. W zerojedynkowej tablicy odtworzonych sygnałów pełny wiersz określa postać odebranego symbolu. Przy jednoczesnej pracy wielu współużytkowników systemu możliwe jest powstawanie więcej niż jednego pełnego wiersza w tablicy sygnałów odtworzonych. Może to powodować błędy decyzyjne przy losowym wyborze symbolu na podstawie różnych pełnych wierszy. Jednak jak wykazano w [3], [4], [5], przy znajomości adresów współużytkowników systemu jest możliwa eliminacja tych pełnych wierszy, które nie są wynikiem przesyłania symbolu do rozpatrywanego użytkownika.

4. WPŁYW ZAKŁÓCEŃ NA JAKOŚĆ SYSTEMU FFH CDMA

Podstawowym problemem w militarnych zastosowaniach systemu FFH z niedeterministycznym wybieraniem nośnych jest określenie podatności tego systemu na zakłócenia celowe, które z założenia mają charakter addytywny.

Rozpatrywany system FFH charakteryzuje się tym, że każdy z odbiorników w sposób ciągły rejestruje sygnały na każdej z częstotliwości hoppingowych. Wobec tego dowolny sygnał odebrany na tych częstotliwościach można uznać za sygnał zakłócający. Jednak rzeczywiste zakłócenie polegające na błędnym podjęciu decyzji o postaci symbolu przez odbiornik powstaje wówczas, gdy w tablicy sygnałów odtworzonych powstaje więcej niż jeden pełny wiersz. Również jako zakłócenie celowe może być traktowany sygnał występujący w paśmie pojedynczej nośnej systemu FFH, a pochodzący od użytkownika (użytkowników) pracującego w innym analogowym lub cyfrowym systemie radiokomunikacyjnym. Przyjęto, że takie zakłócenie może występować co najmniej przez czas trwania symbolu, a jego moc jest stała. Jeżeli zakłócenie będzie dotyczyć pasma wokół jednej nośnej, wówczas będzie traktowane jako pojedynczy sygnał zagłuszający, natomiast w przypadku zakłócania większej liczby nośnych - jako wieloczęstotliwościowy sygnał zagłuszający.

Jak wykazano w publikacjach [4] i [5], prawdopodobieństwo błędu binarnego systemu FFH bez stosowania procedur tzw. eliminacji zakłóceń pochodzących od współużytkowników systemu w obecności zakłóceń wieloczęstotliwościowych jest określone wzorem

$$P_{bRUZ} = \frac{M}{2(M-1)} \sum_{j=1}^{K-1+Z} \frac{j}{j+1} B(j; M-1; P_{RUZ}) \quad (4.1)$$

$$P_{RUZ} = \left[1 - \left(1 - \frac{1}{M} \right)^{K-1+Z} \right]^L \quad (4.2)$$

gdzie:

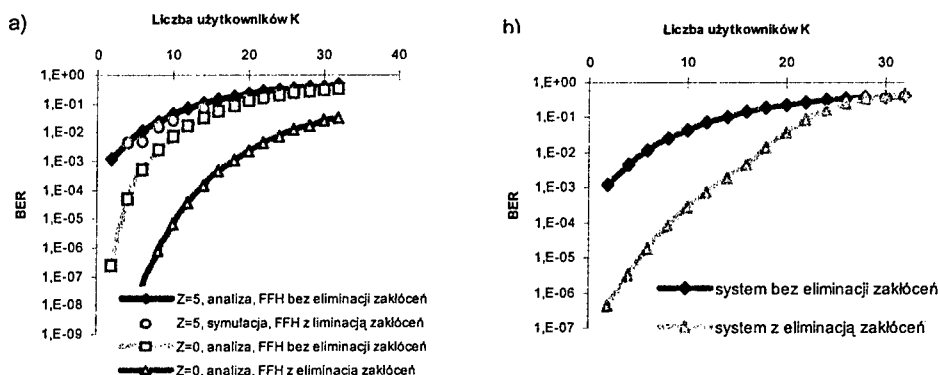
K – liczba aktywnych użytkowników systemu,

$B(, ,)$ – dwumianowy rozkład Bernoulliego.

Zakłócenia celowe, które występują na danej częstotliwości nośnej przez co najmniej czas trwania symbolu powodują taką samą degradację jakości odbioru jak sygnały pochodzące od współużytkowników systemu. Na rys. 1a został przedstawiony wpływ liczby zakłócających częstotliwości nośnych na jakość odbioru.

4.1. Jakość pracy systemu FFH z eliminacją celowych zakłóceń wieloczęstotliwościowych i zakłóceń pochodzących od współużytkowników

Za pomocą symulacji komputerowej dokonano oceny jakości odbioru w systemie FFH w obecności wieloczęstotliwościowych zakłóceń celowych. System FFH wykorzystywał procedury eliminacji zakłóceń pochodzących od współużytkowników. Wyniki badań przedstawione na rys. 1a dowodzą, że metoda eliminacji zakłóceń pochodzących od współużytkowników nie jest skuteczna w obecności zakłóceń celowych [3]. Jakość systemu jest praktycznie taka sama, jak systemu, w którym nie zastosowano żadnych procedur eliminacji zakłóceń.



Rys.1. Prawdopodobieństwo błędu binarnego dla systemu, gdy $M=32$, $L=5$ w obecności zakłóceń wąskopasmowych, liczba zakłóconych nośnych $Z=5$: a) dla systemu z eliminacją zakłóceń tylko od współużytkowników; b) dla systemu z eliminacją zakłóceń od współużytkowników i zakłóceń wieloczęstotliwościowych.

Możliwe jest jednak eliminowanie takich zakłóceń, które występują na danej częstotliwości nośnej przynajmniej przez czas trwania symbolu. Wówczas takie zakłócenia można traktować jak symbol wysłany do użytkownika, którego wszystkie składowe wektora adresu mają jednakowe wartości, np. $\vec{a}_Z = (0, 0, \dots, 0_L)$. Szczegółowy opis metody eliminacji celowych zakłóceń wieloczęstotliwościowych w tym systemie przedstawiono w pracach [4], [5]. Wówczas prawdopodobieństwo powstania błędu binarnego jest dane zależnością

$$P_{bez} = \left(P_{RUZ} + P_{MRUZ} \cdot \sum_{n=2D}^{Q_Z} \frac{n}{Q_Z} \cdot B(n; Q_Z; P_{MRUZ}) \right) \cdot \frac{M}{2(M-1)} \sum_{j=1}^Y \frac{j}{j+1} \cdot B(j; M-1; P_{RUZ}) \quad (4.3)$$

gdzie:

$$EW_Z = \sum_{j=1}^Y j \cdot B(j; M-1; P_{RUZ}) \quad (4.4)$$

$$Y = K-1+Z \quad \text{dla } K-1+Z \leq M \quad (4.5)$$

$$Y = M \quad \text{dla } K-1+Z > M$$

$$P_{MRUZ} = \left[1 - \left(1 - \frac{1}{M} \right)^{Q_Z+Z} \right]^L \quad (4.6)$$

$$D = \lfloor EW_z + 1 \rfloor \quad (4.7)$$

$$Q_z = \lfloor K(EW_z + 1) \rfloor \quad (4.8)$$

Jakość systemu z eliminacją zakłóceń od współużytkowników systemu i celowych zakłóceń wieloczęstotliwościowych, wyznaczoną analitycznie na podstawie powyższych wzorów, przedstawia rys. 1b.

Zakłócenie celowe występujące na jednej z nośnych powoduje taką samą degradację jakości odbioru jak zakłócenie wywołane sygnałami współużytkownika. Eliminacja zakłóceń jest bardziej efektywna przy małej i średniej liczbie aktywnych użytkowników. W przypadku dużej liczby użytkowników, bliskiej liczbie częstotliwości nośnych, procedury eliminacji zakłóceń są nieskuteczne. W powyższej analizie nie uwzględniono szumu w kanale radiowym, który będzie zwiększać degradację jakości odbioru.

4.2. Wpływ szumu na jakość odbioru w systemie FFH

Jako wiadomo, szum termiczny jest nieodłącznym zjawiskiem towarzyszącym przesyłaniu sygnałów w kanale i wpływa degradującą na jakość odtwarzania w odbiorniku. W celu przeprowadzenia analizy wpływu szumu na jakość odbioru w systemie radiokomunikacyjnym założono, że moc szumu dla pasma wokół każdej z częstotliwości nośnych systemu jest jednakowa. Szczegółowa analiza wpływu szumu na jakość systemu FFH zastała przeprowadzona w publikacjach [6] i [7]. Istotnym problemem w tej analizie było określenie prawdopodobieństwa błędu reprezentującego tzw. fałszywy alarm. Jest ono dane wzorem

$$P_f = \int_{\alpha}^{\infty} \frac{2v}{N} \cdot e^{-\left(\frac{v^2}{N}\right)} dv = e^{-\left(\frac{\alpha^2}{N}\right)} = e^{-\left(\alpha'^2\right)} \quad (4.9)$$

gdzie: v – moduł wartości chwilowej szumu,
 N – moc szumu,
 α – wartość progu detekcji,
 $\alpha' = \alpha/\sqrt{N}$ znormalizowany próg detekcji.

Natomiast prawdopodobieństwo błędu oznaczającego brak alarmu, pomimo obecności sygnału użytecznego w sygnale odebranym, przyjmuje postać

$$P_d = \int_0^{\alpha} \frac{2v}{N} \cdot e^{-\left(\frac{v^2+A^2}{N}\right)} \cdot I_0\left(\frac{2vA}{N}\right) dv \quad (4.10)$$

Wobec tego prawdopodobieństwo błędu binarnego w systemie FFH bez eliminacji zakłóceń własnych systemu w obecności szumu jest określone zależnością

$$R_{bn} = \frac{M}{2(M-1)} \left\{ \frac{M-1}{M} [1 - (1-P_d)^L] + \sum_{j=1}^{M-1} \frac{j}{j+1} B(j, M-1, P_d^L(P_f)) \right\} \quad (4.11)$$

gdzie P_d oznacza prawdopodobieństwo detekcji sygnału wskutek emisji pochodzącej od współużytkowników systemu lub szumu w kanale, (wzory wyznaczające wartość tego parametru podane są w publikacjach [6] i [7]).

Natomiast jakość systemu FFH w obecności szumu i z eliminacją zakłóceń pochodzących od współużytkowników można określić za pomocą wzoru

$$R_{bn} = \frac{M}{2(M-1)} \left\{ \frac{M-1}{M} [1 - (1-P_d)^L] + (P_{RU} + P_{CB} + P_{DU}) \sum_{j=1}^{M-1} \frac{j}{j+1} B(j, M-1, P_d^L(P_f)) \right\} \quad (4.12)$$

gdzie $\frac{M-1}{M}[1-(1-P_d)^M]$ oznacza prawdopodobieństwo nieodebrania przynajmniej jednego chipu z ciągu chipów tworzących symbol przeznaczony dla określonego użytkownika i podjęcie przez człon decyzyjny błędnej decyzji o postaci symbolu. Aby eliminacja zakłóceń była również możliwa w innych istotnych dla procesu eliminacji zakłóceń tablicach odtworzonych sygnałów (tworzonych dla adresów innych odbiorników) nie mogą występować pełne wiersze z przynajmniej jednym chipem wywołanym wyłącznie wskutek występowania szumu. Stąd we wzorze (4.12) P_{CB} oznacza prawdopodobieństwo utworzenia pełnego wiersza, w którym pojawi się chip wywołany przez szum, a P_{DB} to prawdopodobieństwo powstania takich tablic odtworzonych sygnałów, w których nie jest możliwa eliminacja błędnych wierszy.

Przedstawione powyżej rozważania dotyczyły przypadku, gdy sygnał odbierany był tylko jednokrotnie próbkowany w czasie trwania chipu. Poprawę jakości pracy systemu można uzyskać poprzez zastosowanie wielokrotnego próbkowania, gdyż q -krotne próbkowanie pozwoli na \sqrt{q} -krotne zmniejszenie wpływu szumu.

4.3. Wpływ szumowych sygnałów zagłuszających na jakość pracy systemu FFH

W przypadku, gdy celowe zakłócenie ma postać zaporowego, szumowego sygnału zagłuszającego, występuje ono z jednakową mocą w paśmie każdej z częstotliwości nośnych wykorzystywanych przez system FFH. Wówczas jakość pracy systemu może być określona podobnie jak dla zwykłego tła szumowego za pomocą wzorów (4.11) – dla systemu bez eliminacji zakłóceń i (4.12) – dla systemu z eliminacją zakłóceń od współużytkowników. W rzeczywistości zakłócanie szumem całego pasma systemu jest mało prawdopodobne, ze względu na ogromną moc niezbędną do uzyskania pożądanego skutku. W przypadku szumowych sygnałów zagłuszających należy spodziewać się, że tylko część pasma wykorzystywanego przez system FFH będzie objęta tego rodzaju zakłóceniem. Oczywiście możliwy jest scenariusz, że w pewnym podzakresie pasma systemu będzie jednocześnie oddziaływać pewna liczba U generatorów szumowych sygnałów zagłuszających, a każdy z nich będzie zagłuszał pasmo wokół m_u ($u=1,2,\dots,U$) częstotliwości nośnych. Jeżeli pasma sygnałów zagłuszających poszczególnych generatorów będą rozdzielne i nie wszystkie pasma wokół nośnych zostaną zakłócone, w systemie będzie występować $U+1$ wartości¹ mocy szumu. Wobec tego średnie prawdopodobieństwo błędu oznaczającego fałszywy alarm, czyli detekcję sygnału pomimo braku sygnału użytecznego, będzie dane zależnością

$$P_{fu} = \frac{1}{M} \sum_{u=1}^{U+1} m_u P_{fu} \quad (4.13)$$

gdzie:

$$P_{fu} = \int_{\alpha}^{\infty} \frac{2v}{N_u} \cdot e^{-\left(\frac{v^2}{N_u}\right)} dv = e^{-\left(\frac{\alpha^2}{N_u}\right)} \quad (4.14)$$

oznacza prawdopodobieństwo fałszywego alarmu w u -tym paśmie, przy czym N_u jest mocą szumu zagłuszającego w u -tym paśmie, a m_u oznacza liczbę nośnych objętych szumem o mocy N_u . Natomiast średnie prawdopodobieństwo błędu oznaczającego brak alarmu, pomimo obecności sygnału użytecznego przyjmuje postać

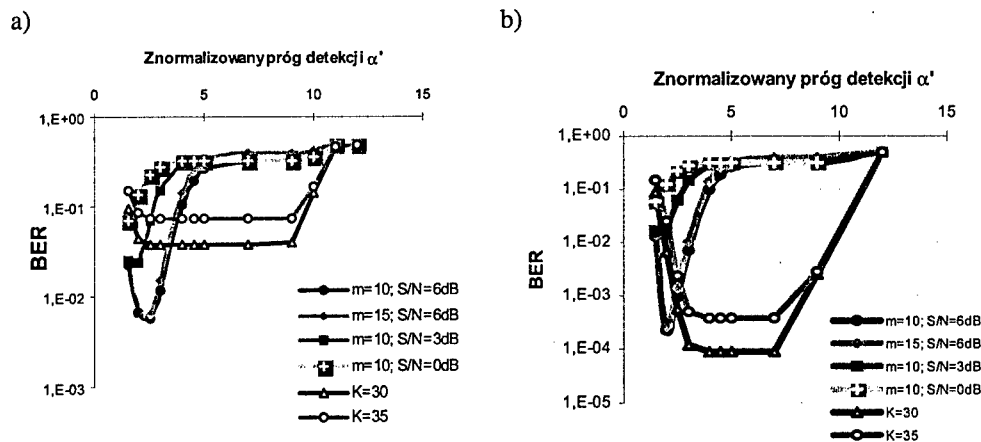
¹ Uwzględniono dodatkowo zbiór tych nośnych, w obrębie których występuje tylko szum naturalny

$$P_{dU} = \frac{1}{M} \sum_{u=1}^{U+1} m_u P_{du} \quad (4.15)$$

gdzie

$$P_{du} = \int_0^{\alpha} \frac{2v}{N_u} \cdot e^{-\left(\frac{v^2 + \lambda^2}{N_u}\right)} \cdot I_0\left(\frac{2v\lambda}{N_u}\right) dv \quad (4.16)$$

oznacza prawdopodobieństwo braku alarmu w u -tym paśmie. Po uwzględnieniu uśrednionych wartości P_{dU} i P_{fU} zamiast P_d i P_f we wzorach (4.11) i (4.12) można wyznaczyć jakość pracy systemu FFH w przypadku zagłuszania szumem części pasma systemu.



Rys. 2. Jakość pracy systemu FFH przy różnej liczbie zakłócających nośnych i różnej mocy sygnału zakłócającego. Parametry systemu: $M=64$, $L=6$, $q=5$, $K=20$, $S/N=14$ dB dla niezagłuszanych nośnych: a) system bez eliminacji zakłóceń; b) system z eliminacją zakłóceń.

Jakość ta dla systemu FFH o parametrach: $M=64$, $L=6$, $q=5$, dla dwudziestu aktywnych użytkowników, przy stosunku mocy sygnału do szumu wynoszącym 14dB dla niezagłuszanych nośnych jest przedstawiona na rys. 2. Wszystkie zagłuszane nośne, których liczba wynosi m , znajdują się pod działaniem tego samego sygnału zagłuszającego. Na rys. 2 porównano jakość pracy systemu przy zakłócaniu 10 oraz 15 częstotliwości nośnych szumem o różnych mocach względem mocy sygnału użytecznego. W celu porównania jakości pracy systemu FFH przy zagłuszaniu szumem i braku tego typu zakłócenia, wyznaczono binarną stopę błędów dla systemu FFH przy liczbie aktywnych użytkowników stanowiącej sumaryczną wartość zagłuszanych szumem nośnych i aktywnych użytkowników. Jakość pracy systemu FFH z eliminacją zakłóceń przy zwiększonej liczbie użytkowników i braku zagłuszania szumem jest nie gorsza niż w przypadku zagłuszania szumem, a nawet jest zdecydowanie lepsza w porównaniu z przypadkami zagłuszania szumem o bardzo dużej mocy. Ponadto przedział wartości znormalizowanego progu detekcji jest większy niż w przypadku zagłuszania szumem, wartość tego progu nie musi być precyzyjnie dobierana. W takim przypadku, w celu przeciwdziałania zakłóceniom sygnałami szumopodobnymi, należy na każdej częstotliwości nośnej prowadzić ocenę mocy szumu. Gdy, na danej nośnej moc szumu będzie porównywalna z mocą sygnału użytecznego, to dla tej nośnej w tablicy sygnałów odebranych należy podjąć decyzję o detekcji wszystkich chipów przez cały czas trwania symbolu. Wówczas, jakość pracy

systemu będzie taka, jak w przypadku zakłócania takiej samej liczby nośnych sygnałami jednotonowymi.

5. WNIOSKI

Proponowany system radiokomunikacyjny może być szczególnie atrakcyjny dla morskich zastosowań militarnych. Może on ze względu na krótki czas trwania emisji energii na poszczególnych nośnych stwarzać problemy przeciwnikowi w lokalizowaniu źródeł sygnału. A ponadto różni użytkownicy wykorzystujący te same częstotliwości nośne wnoszą dodatkową dezorientację w systemach rozpoznania radiowego przeciwnika. Znaczna odporność tego systemu na zakłócenia wieloczęstotliwościowe i szumowe przy zastosowaniu wyżej opisanych metod powoduje, że praktycznie jakość systemu zależy od liczby zakłócanych nośnych, a wzrost mocy zakłóceń nie powoduje dalszej degradacji jakości systemu. System FFH, jako system szerokopasmowy, dokonujący rozpraszania widma sygnału użytecznego, przeciwdziała także zanikom występującym w kanale radiokomunikacyjnym. Należy oczekiwać, że system FFH CDMA z niedeterministycznym wybieraniem nośnych przy odpowiednim jego zorganizowaniu pozwoli zachować okrętom w większym stopniu skrytość działania, niż inne systemy radiokomunikacyjne, przy zapewnieniu wysokiej jakości odbioru danych zarówno na falach ultrakrótkich jak i krótkich.

BIBLIOGRAFIA

- [1] Einarsson G.: *Address Assignment for a Time-Frequency-Coded Spread-Spectrum System*, The Bell System Technical Journal vol. 59, nr 7, 1980, str. 1241-1255.
- [2] Goodman D. J., Henry P. S., Prabhu V. K.: *Frequency-Hopped Multilevel FSK for Mobile Radio*, The Bell System Technical Journal vol. 59, nr 7, 1980, str. 1257-1275.
- [3] Rutkowski D., Studański R.: *Performance of FFH CDMA system with mobile radio channel*, NATO Regional Conference on Military Communications and Information Systems 2001, Zegrze 2001, str. 100-106.
- [4] Studański R.: *Wpływ zakłóceń wąskopasmowych na jakość transmisji danych w systemie FH CDMA z szybkim wybieraniem nośnych*; VII Krajowa Konferencja Naukowo-Dydaktyczna "Automatyzacja i eksploatacja systemów sterowania", Gdynia 1999, str. 341-348.
- [5] Rutkowski D., Studański R.: *Performance of FFH CDMA system with the rejection of own and intentional interference*, Regional Conference on Military Communication and Information Systems, Zegrze 2000, tom 3, str. 155-160.
- [6] Rutkowski D., Studański R.: *Performance of FFH CDMA system with mobile radio channel*, NATO Regional Conference on Military Communications and Information Systems 2001, Zegrze 2001, str. 100-106.
- [7] Rutkowski D., Studański R.: *Jakość systemu FFH CDMA w kanale z zanikami i szumem*, VIII Krajowa Konferencja Naukowo-Dydaktyczna "Automatyzacja i eksploatacja systemów sterowania i łączności", Gdynia 2001, str. 1159-168.

TRENDS IN THE DEVELOPMENT DATA TRANSMISSION RADIO COMMUNICATION SYSTEMS FOR THE NAVY

Summary

The performance of signal transmission in the radio communication system based on fast frequency hopping with nondeterministic carrier selection has been described. The effects of rejection methods of jamming on system performance have been shown.

Marek Wesółowski, Mariusz Zieliński

Akademia Marynarki Wojennej

ZASTOSOWANIE KOMPUTEROWEJ REJESTRACJI MOWY W MONITORINGU RADIOWYM NA MORZU

Streszczenie

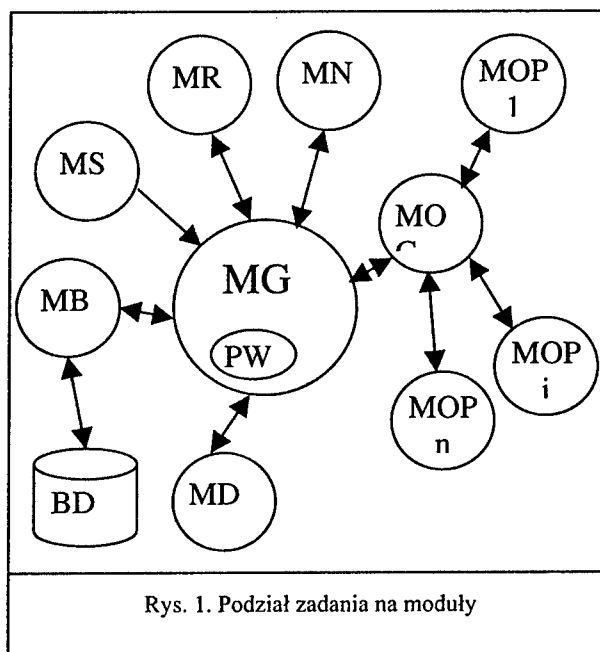
W artykule opisano stanowisko monitoringu radiowego opartego o odbiornik i namiernik firmy Rhode&Schwarz, jako elementu systemu rozpoznania elektronicznego. Stanowisko zostało poszerzone o komputerową rejestrację emisji radiowych. W Akademii Marynarki Wojennej trwają prace nad wyborem metody zapisu dźwięku. Jednocześnie testowane są możliwości sterowania kilku odbiorników połączonych w sieć lokalną pozwalającą na jednoczesne przeszukiwanie różnych pasm lub kanałów częstotliwości. Rozwiązanie oparto o system czasu rzeczywistego QNX oraz język C++. Zastosowano podział zadania na osobne i niezależne moduły programowe komunikujące się poprzez segmenty pamięci wspólnej.

1. WSTĘP

W Akademii Marynarki Wojennej opracowano zintegrowane stanowisko rozpoznania radiowego dla potrzeb marynarki Wojennej RP. Ze względu na długoletnie pozytywne doświadczenia ze sprzętem firmy Rohde&Schwarz oraz parametry i możliwości nowych konstrukcji tej firmy stanowisko oparto o odbiornik ESMB i namiernik DDF 190. Aczkolwiek brano również pod uwagę sprzęt firmy Thompson, o porównywalnych możliwościach. Jako uzupełnienie stanowiska zastosowano urządzenie GPS i analogowy rejestrator emisji radiowych, pozwalający na rejestrację wybranych emisji na kasetach magnetofonowych. Zestaw ten umożliwia realizację podstawowych funkcji:

- nasłuch emisji radiowych dla wybranej częstotliwości (wraz z rejestracją);
- śledzenie i kontrolę łączności w wybranych pasmach częstotliwości (z ustaloną wielkością kroku);
- śledzenie i kontrolę łączności w wybranych kanałach częstotliwości (każdy kanał może mieć inne parametry);
- określenie kierunku, (względem północy) na wykrytą emisję.
- obsługa bazy danych;
- przygotowanie i wywoływanie algorytmów przeszukiwania.

Uwzględniając potrzebę szybkiej, automatycznej obsługi oraz wydajność i niezawodność pracy, parametry szczególnie ważne w zastosowaniach wojskowych i ratowniczych, zastosowano wielozadaniowy system czasu rzeczywistego dla sieci lokalnych QNX. System ten daje duże możliwości przetwarzania rozproszonego, pozwala na wszechstronną kontrolę uruchomianych procesów oraz umożliwia komunikację i synchronizację między nimi. Aplikacja została wykonana w języku C++. Stanowisko obsługuje tylko jeden operator. Ze względu na fakt, iż główne elementy muszą pracować niezależnie od siebie w tym samym czasie ale jednocześnie muszą być zsynchronizowane, zadanie zostało podzielone na oddzielne moduły, którymi zarządza moduł główny (rys 1. – docelowy podział zadania na moduły).



gdzie: MG – moduł główny,
 MB – moduł bazy,
 MS – moduł GPS,
 MR – moduł komputerowego rejestratora wykrytych sygnałów,
 MN – moduł namiernika,
 MOG – moduł odbiornika głównego,
 MOP1 – moduł odbiornika pomocniczego (1),
 MOPi – moduł odbiornika pomocniczego (i-ty),
 MOPn – moduł odbiornika pomocniczego (n-ty),
 MD – moduł wysyłania meldunków,
 BD – baza danych SQL,
 PW – pamięć wspólna.

Tak zestawione moduły wymagają odpowiedniego, wydajnego i szybkiego sposobu komunikacji oraz kontroli wykonywanych zadań. System QNX daje trzy możliwości komunikacji:

- synchroniczny przekaz wiadomości;
- sygnalizację zdarzeń;
- asynchroniczny przekaz depozytów.

Najbardziej rozpowszechniony jest pierwszy sposób, niestety ma on bardzo istotną wadę, polegającą na zawieszaniu procesu głównego na czas przekazywania komunikatu. W opracowanej aplikacji zastosowano niestandardowy sposób synchronizacji modułów polegający na połączeniu możliwości synchronicznego przekazu wiadomości oraz sygnalizacji zdarzeń z zastosowaniem segmentów pamięci wspólnej (ang. *shared memory object*). Pamięć wspólna jest raczej rzadko stosowanym elementem w procesie tworzenia oprogramowania. Jednak przy rozbiciu zadania na oddzielne moduły daje ona przyspieszenie pracy i możliwość niezależnej pracy wszystkich modułów dzięki przekazywaniu komunikatów, poleceń i parametrów bez konieczności zawieszania pracy na czas połączenia. Rozbicie zadania na oddzielne moduły programowe pozwala znacznie przyspieszyć wykonanie całego procesu poprzez współbieżną obsługę urządzeń oraz współbieżne wykonywanie modułów niezależnych. Przy założeniu pracy w rozproszonym środowisku czasu rzeczywistego otrzymujemy możliwość dynamicznego przydziału modułów do komputera posiadającego wolne moce obliczeniowe (mało obciążonego).

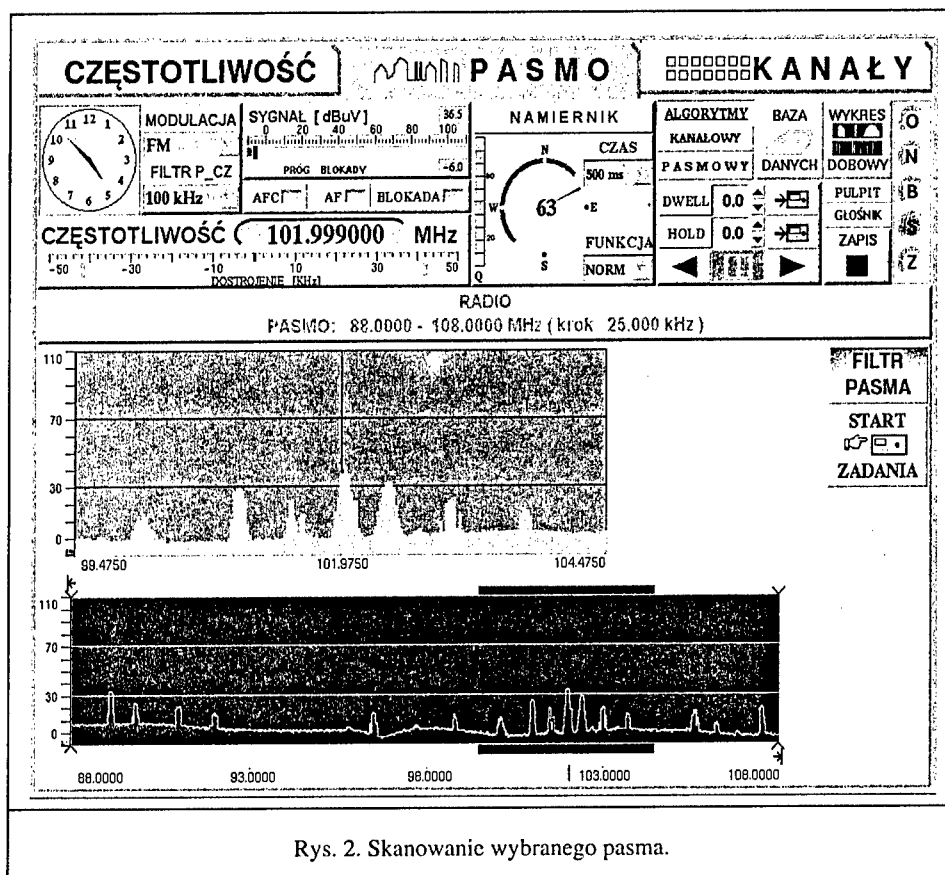
2. DZIAŁANIE APLIKACJI

Ekran sterowania urządzeniami podzielono na dwie części: górną, wspólną dla wszystkich funkcji oraz dolną obrazującą wykryte emisje w sposób właściwy dla każdej z funkcji. Funkcja nasłuchu wybranej częstotliwości umożliwia rejestrację amplitudy wykrytego sygnału w czasie (do 1 godziny). Funkcja przeszukiwania wybranego pasma częstotliwości z zadaniem krokiem (ang. *frequency scanning*) wizualizuje widmo wykrytego sygnału w postaci dwóch wykresów (Rys. 2): przybliżonego (dolny) i dokładnego (górny). Operator w dowolnym momencie może zmienić wykres dokładny i obejrzeć każdy podzakres skanowanego pasma. Biorąc pod uwagę szybkość zmiany sygnałów wprowadzono rozróżnienie wykrytych emisji innymi kolorami:

- zielony – wykrycie sygnału podczas ostatniej pętli skanowania,
- żółty – wykrycie sygnału w przedostatniej pętli skanowania (brak potwierdzenia w pętli ostatniej),
- szary – wykrycie sygnału w ostatnich dziesięciu pętlach skanowania (brak potwierdzenia w ostatniej i przedostatniej).

Wprowadzono możliwość bardzo szybkiego zatrzymania skanowania na wykrytej emisji. Funkcja przeszukiwania wybranych kanałów częstotliwości (ang. *memory scanning*) pozwala na zaprogramowanie do 100-u kanałów jednocześnie, każdy może mieć inne parametry. Dla wszystkich kanałów zapamiętywany jest wykres amplitudy sygnału w czasie (do 1 godziny). Wprowadzono możliwość porównywania do trzech wykresów jednocześnie, dla trzech dowolnie wybranych kanałów. Operator jest informowany o wykryciu emisji w kanałach lampką ostrzegawczą. Dowolny kanał można włączyć lub wyłączyć z procesu skanowania jednym ruchem myszki (maskowanie kanału). Podobnie jak w funkcji pasmowej wprowadzono możliwość bardzo szybkiego zatrzymania skanowania na wykrytej emisji w dowolnym kanale. W celu automatycznego i szybkiego

programowania urządzeń (przede wszystkim odbiornika), wprowadzono moduł opracowywania i wyboru algorytmów śledzenia pasmowego i kanałowego. Jednocześnie może być zapamiętanych do 200-u algorytmów pasmowych (każdy o innych parametrach: zakres pasma, modulacja poziom blokady szumów) i do 200-u algorytmów kanałowych (każdy do 100-u różnych kanałów). Wszystkie wykryte emisje są automatycznie (bez udziału operatora) zapisywane do bazy danych i w dowolnym momencie można dokonać przekrojowej analizy zapisanych sygnałów, stosując dowolne kryteria wyszukiwania. Informacje można uzyskać w postaci wykresu wykrytych emisji dla pełnego zakresu odbiornika. Wykres górny jest wykresem przybliżonym (z dokładnością do 1 MHz), natomiast wykres dolny umożliwia analizę ze zwiększoną dokładnością wybranego podzakresu (wprowadzono ZOOM). Wyłączenie ciekłokrystalicznego pulpitu sterowania odbiornikiem znacznie przyspiesza skanowanie pasma lub wybranych kanałów częstotliwości. Ze względu na brak miejsca na ekranie komputera część informacji jest ukryta i wywoływana odpowiednim klawiszem. Np.: jeśli naciśniemy myszką w tarczę zegara analogowego pojawi się ukryta pod tarczą informacja o aktualnej pozycji własnej.



Rys. 2. Skanowanie wybranego pasma.

3. REJESTRATOR MOWY

Ze względu na ograniczenia finansowe, tak zestawione stanowisko ma kilka istotnych wad. Zastosowano jeden z dostępnych na rynku prostych rejestratorów rozmów zapisujący dźwięk na kasie magnetofonowej. Podczas eksploatacji w warunkach rzeczywistych wybrany rejestrator okazał się urządzeniem awaryjnym. Również rejestracja na takim nośniku w warunkach podwyższonej wilgotności (warunki morskie) nie jest rozwiązaniem optymalnym. Zastosowany rejestrator pozwala tylko i wyłącznie na rejestrację, natomiast do odtwarzania potrzebny jest specjalny odtwarzacz. Z powyższych powodów aktualnie w AMW trwają prace nad zastosowaniem komputerowego rejestratora mowy. Aczkolwiek na rynku są dostępne gotowe produkty, są one drogie i nie spełniają wszystkich naszych wymagań, dlatego postanowiono zrealizować własne rozwiązanie. Wykorzystano wyjście audio odbiornika ESMB i połączono z wejściem karty dźwiękowej komputera. Szybki rozwój pojemności dysków sztywnych pozwala na rejestrację od kilku do kilkunastu tysięcy godzin emisji radiowej na pojedynczym nośniku, oraz łatwą archiwizację na nośnikach CD-R/CD-RW. W celu zwiększenia ilości rejestrowanych godzin podczas prób stosowano kompresję dźwięku. W poniższych tabelkach przedstawiono porównanie czasów rejestracji (w godzinach) dla kompresji GSM 6.10 oraz zapisu bez kompresji PCM dla ośmiu bitów.

Tabela 1

Czasy rejestracji dla zapisu z kompresją GSM

Częstotliwość / kHz /	Pojemność dysku / GB /		
	40	80	120
8,000	13 605	27 210	40 816
11,025	9 803	19 607	29 411
22,050	4 866	9 732	14 598
44,100	2 469	4 938	7 407

Tabela 2

Czasy rejestracji dla zapisu PCM bez kompresji

Częstotliwość / kHz /	Pojemność dysku / GB /		
	40	80	120
8,000	1 388	2 777	4 166
11,025	1 010	2 020	3 030
12,000	900	1 801	2 702
16,000	694	1 388	2 083
22,050	505	1 010	1 515
24,000	453	907	1 360
32,000	347	694	1 041
44,100	245	490	735
48,000	226	453	680

Jak można wywnioskować z powyższych tabel zastosowanie kompresji pozwala znacznie zwiększyć czas rejestrowanych danych. Dzieje się to oczywiście kosztem jakości zapisu, ale tam gdzie nie jest wymagana najwyższa jakość oraz gdy dźwięk nie będzie poddawany dalszej analizie laboratoryjnej zastosowanie kompresji GSM 6.10 z częstotliwością próbkowania 44,1 kHz daje zadowalające rezultaty i co bardzo ważne pozwala na zapis do 308 dni ciągłej rejestracji na dysku o pojemności 120 GB. Natomiast w zastosowaniach w których zapisane dane poddawane będą analizie laboratoryjnej, co oznacza że parametry dźwięku mogą być bardzo istotne, należy zastosować zapis PCM bez kompresji. Biorąc pod uwagę parametry odbiornika ESMB wystarczający będzie format zapisu PCM z częstotliwością próbkowania 16 kHz. Umożliwi to zarejestrowanie do 86 dni ciągłej rejestracji na dysku o pojemności 120 GB. Zastosowanie technologii DirectX pozwala na szybką rejestrację (lub odtworzenie nagranych danych) bez angażowania mocy procesora, co nie wstrzymuje pracy innych procesów.

4. DODATKOWE ODBIORNIKI

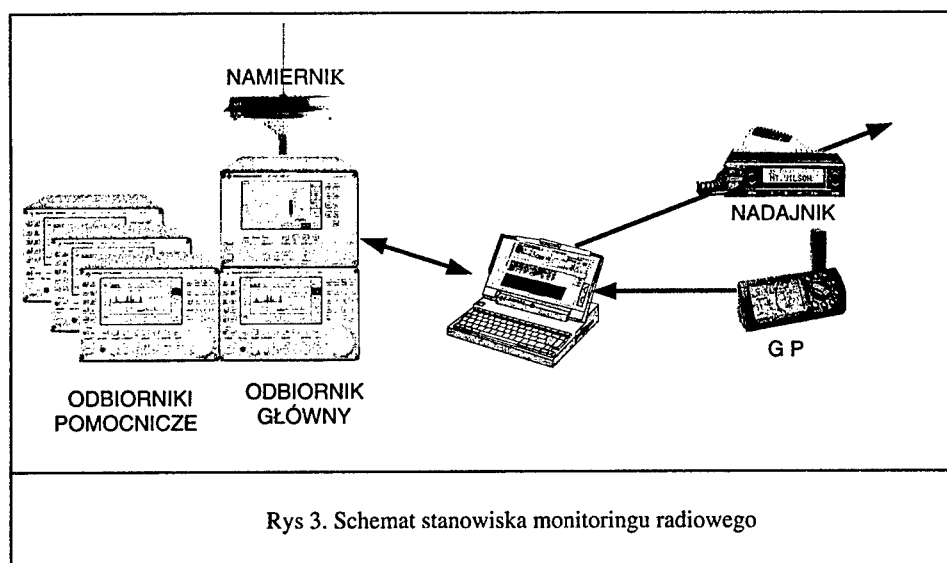
Jeszcze większą wadą okazała się obecność tylko jednego odbiornika, co uniemożliwia jednoczesny nasłuch wybranej częstotliwości oraz śledzenie wybranego pasma lub kanałów częstotliwości. Aktualnie trwają prace z zastosowaniem dodatkowego odbiornika, a jeśli możliwości finansowe pozwolą nawet kilku odbiorników połączonych w sieć pozwalającą na jednoczesny nasłuch kilku częstotliwości. Dotychczas stosowany odbiornik wyposażony jest w łącznie szeregowo wprowadzając istotne ograniczenie szybkości skanowania pasma i kanałów.

W nowej wersji zastosowane będą odbiorniki z łączem sieciowym pozwalającym na bardzo szybką transmisję pomiędzy komputerem i odbiornikami oraz umożliwiające włączenie ich w sieć lokalną i sterowanie z jednego komputera. Dodatkowa korzyść wynika z zastosowania odbiorników ESMB o paśmie do 3,0 GHz z możliwością skanowania cyfrowego (opcja DIGI Scan) co umożliwi skanowanie wybranego pasma częstotliwości z bardzo dużą szybkością do 3000 MHz/s (przy filtrze IF 300 kHz). Dla porównania standardowa funkcja przeszukiwania pasma (Fscan) umożliwia skanowanie z szybkością do 250 kanałów/s (przy filtrze IF 300 kHz), natomiast funkcja przeszukiwania wybranych kanałów częstotliwości (Mscan) pozwala przeglądać do 200 kanałów/s (przy filtrze IF 300 kHz).

5. ZAKOŃCZENIE

Zastosowanie komputerowej rejestracji pozwala na rejestrację dodatkowych parametrów (czas rozpoczęcia i zakończenia zapisu; parametrów sygnału: modulacja, częstotliwość, zastosowane filtry). Tak zarejestrowane emisje można bardzo szybko odczytać i wszechstronnie przeanalizować w odpowiednim laboratorium. Użycie kilku odbiorników sterowanych z jednego komputera pozwoli na jednoczesny nasłuch różnych częstotliwości. Natomiast odbiorniki nowszej generacji znacznie przyspieszają cykl skanowania wybranego pasma. Zestaw składający się z przenośnego komputera, odbiornika, namiernika oraz przewoźnej anteny można zastosować również jako ruchomy punkt namierzania. Połączenie co najmniej dwóch takich punktów drogą radiową pozwoli określić pozycję wykrytego obiektu. Podobnie jak w poprzednich wersjach również teraz zwracano uwagę na prostotę i szybkość obsługi. stanowiska. Obsługa aplikacji jest bardzo prosta wręcz intuicyjna, co zwiększa możliwości jej wykorzystania w innych zautomatyzowanych

systemach kontroli, rozpoznania lub dowodzenia, jako dostarczyciel informacji z pasma radiowego UKF oraz w systemach ratownictwa morskiego i lotniczego. Aplikacja została w całości wykonana w języku polskim. Na rysunku 3 przedstawiono docelowy schemat stanowiska monitoringu radiowego.



BIBLIOGRAFIA

- [1] Sacha K. „QNX - system operacyjny”. X-serwis. Warszawa 1995.
- [2] QNX OS System Architecture, QNX Software Systems, 1993r.
- [3] Zieliński M., Wesołowski M. „Software for the UHF Signals Intelligence Workstation”. Materiały Regional Conference on Military Communication and Information Systems '99 CIS Solutions for an Enlarged NATO.
- [4] Balicki J., Wesołowski M., Zieliński M. „Maksymalizacja sprawności wspomagania zadań na okręcie za pomocą wyznaczania rozdziału modułów programów w rozproszonym systemie operacyjnym czasu rzeczywistego.” Materiały VII Krajowej Konferencji Naukowo-Dydaktycznej. Gdynia 13-15 października 1999r.
- [5] Wesołowski M., Zieliński M. „The Post of Recognition the Ukf Band”. Materiały Regional Conference on Military Communication and Information Systems '2000 Partnership for CIS Interoperability.

USE IN RADIOMONITORING ON SEA THE COMPUTER REGISTRATION OF SPEECH

Summary

In paper was described the radiomonitoring post including the receiver and direction finder of the Rhode&Schwarc firm, as element of the recognition system. The Post was extended on the computer registration of the radio-emission. In Naval University of Gdynia carry on works over the choice of method of recording of sound. Simultaneously was tested possibility of the steering several receivers working in local net permitting on simultaneous searching of different bands or channels or frequency. Solution was based on real time operating system QNX with C++. There was used partition of exercise on separate and independent programmatic modules communicating across segments of common memory.

Arnold Adamczyk

Katedra Metrologii i Systemów Elektronicznych, Politechnika Gdańska

THE IMPLEMENTATION OF A MICROSERVER

The following paper presented the technical solutions of the microserver with less construction complexity and reduction of a power consumption. Proposed reduction of functionality TCP/IP protocols enable the implementation of correct operated client-server conversation for microcontrollers with limited computational power.

1. OVERVIEW

New devices and applications will exploit the expansion of global communications. These systems with embedded network interfaces will perform in offices, homes, hospitals and factories. The obvious e-commerce applications include network buying and payment systems integrated with home devices [1], homecare accessories connected to far hospital laboratory [2] and systems for remote diagnosis of devices [3, 4]. For those applications are to create a miniature Web server that can monitor and control real world by lines connected to a microcontroller (small single-chip computer). Here are presented the techniques that should be used to minimize the network protocols complexity. From the top level (HTTP – Fig. 1) down, across TCP and IP protocols to the physical layer (Ethernet), are discussed methods to reduce functionality and memory consumption by the miniature Web server.

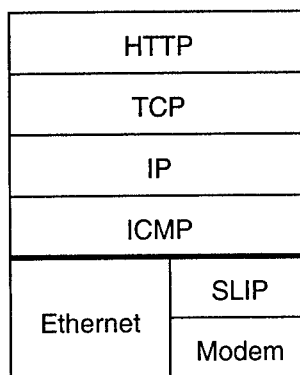


Fig. 1. Web server stack

2. HTTP MICRO-IMPLEMENTATION

The Hypertext Transform Protocol (HTTP) defines a request-response procedure for getting files from a Web server. The client application (Web browser) sends a request to the server in the form of a multi-line text (Fig. 2). The first line specifies command GET or POST and arguments. The most usual command is GET, followed by a filename to be obtained and a protocol identifier. Following lines contain parameters for configuring a browser.

```

GET /webpage.html HTTP/1.0
User-Agent: Mozilla/4.5 [pl] (Win98; I)
Pragma: no-cache
Host: user827.eti.pg.gda.pl
Accept: image/gif, image/jpeg, image/pjpeg, image/png,
Accept-Encoding: gzip
--blank line--

```

Fig. 2. Example of HTTP request.

In the micro-implementation usage of memory and computational complexity could be reduced to a minimum. Ignore firstly all configuration parameters. Only part of first line of request must be processed by microcontroller. The version identifier "HTTP/1.x" isn't needed. Microserver use commands common for all versions of the protocol. In result, only command identifier and filename (with parameters string) are required by the parsing procedure. For more simplicity in practice, only GET command may be implemented.

The below URL-request:

http://www.mymicroserver.com/filename?param1=value1¶m2=value2 (1)

is seen by the procedure as follows string:

1.1. GET /filename?param1=value1¶m2=value2 HTTP/1.0 (2)

The request (1) may be prepared as submission FORM-request and placed in a HTML document (Fig. 3).

```

<FORM method="GET" action=" http://www.mymicroserver.com/filename">
Param1: <INPUT type="text" name="param1" value="value1"><br>
Param2: <INPUT type="text" name="param2" value="value2"><br>
<INPUT type="submit" value="Submit form">
</FORM>

```

Fig. 3. The submission form example.

The controls "param1" and "param2" are editable of course. The values "value1" and "value2" inserted in the form are set as predefined data and may be changed by a user. Web browser into the same string (2) transforms the URL-request and FORM-request. The simple example show, in which one shape form data can received a microserver. So, it will suffice to parse received string (2) for parameters localizing and moving them to memory. But, the procedure must control the request length. If it doesn't detect a space character in string (after filename), then it should signal error. The error message could take following form of standard response:

<pre> HTTP/1.0 414 Request-URL Too Large Content-type: text/html --blank line-- <html><body> <h1>414 Request-URL Too Large</h1> </body></html> </pre>	<div style="font-size: 3em; vertical-align: middle; padding: 0 10px;">}</div> <div>HTTP header</div> <div style="font-size: 3em; vertical-align: middle; padding: 0 10px;">}</div> <div>HTTP body</div>
---	---

Fig. 4. Example of the failure response with error message.

The response consists of an HTTP header and the body separated by a single blank line (Fig 4 and 5). In the micro-implementation, the header should contain the protocol identifier, the status code (success or failure) and the content-type identifier. The (HTML document, plain text, GIF or JPEG graphics).

<pre> HTTP/1.0 200 OK Content-type: text/html --blank line-- <html> <head>My Homepage</head> <body>Page from Microserver</body> </html> </pre>	<div style="font-size: 3em; vertical-align: middle; padding: 0 10px;">}</div> <div>HTTP header</div> <div style="font-size: 3em; vertical-align: middle; padding: 0 10px;">}</div> <div>HTTP body</div>
--	---

Fig. 5. Example of the success response with simple HTML document.

For more simplicity, all files loaded in memory may consist at the beginning of suitable headers. Then the response procedure doesn't need to create the response string.

3. TCP MICRO-IMPLEMENTATION

The Transmission Control Protocol (TCP) provides a reliable connection between client and server. The mechanism of logical conversation between two applications on the network is named as "sockets interface".

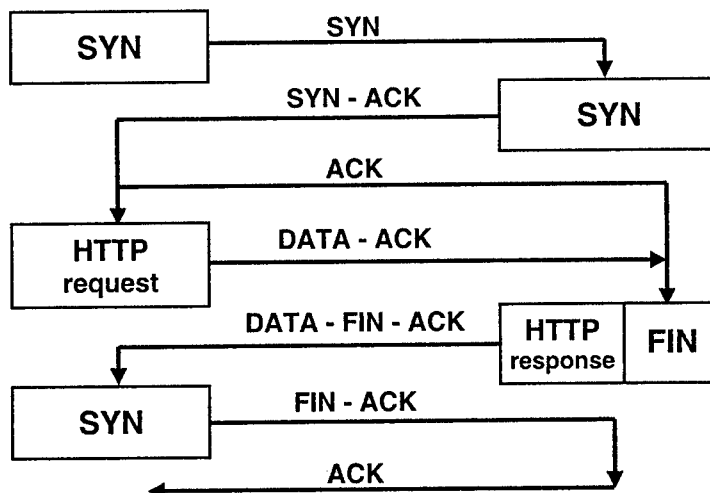


Fig. 5. One-segment TCP conversation.

In the WWW service, a Web-browser is an application on the client side, and the microserver with HTTP procedures (section 2) is an application on the server side. All servers and clients' hosts in the network are identified by the IP addresses. However, network services are identified by the port numbers and in composition with IP's are used to form the sockets. To well know ports belongs, inter alia: HTTP/WWW (80), FTP (20-21), SMTP (25), POP3 (110), NNTP (119), Telnet (23). The microserver provide only the WWW service and then need to operate the port "80".

The TCP conversation process, the starting and stopping of a transaction and the transferring of a data in a TCP segments are identified by a TCP sequence number, which refers to its position in sequence space [5]. The start (identified by SYN) and end (identified by FIN) of a transaction can be seen as steer flags in this space (Fig. 2). The sequencing process is to complicate for use in microcontrollers. A client and server use sequence numbers to place their transmitted segments in the outgoing space and send acknowledgment numbers to confirm the right reception in the incoming sequence space.

The micro-implementation of the TCP protocol is based on the idea of shortening HTML pages and graphics components into the size of a single TCP segment [6]. The one convention of the TCP/IP stack adaptation use the Serial Line Internet Protocol (SLIP [7]) to connect with a physical layer by a serial interface [6, 8]. The maximum SLIP frame size has been set at 1006 bytes. The frame contains the TCP and IP headers; so the maximum data segment size have 966 bytes. The second adaptation is based on a system-bus interface. In this model, to connecting with physical layer don't need use the SLIP protocol [9]. One and only limitation come from the Ethernet frame, which may reach 1518 bytes. After subtracting headers, the segment size may expand to 1460 bytes.

The one-segment strategy decreases the requirement of computational power. The closing FIN adjoined to the page data solidifies this conversation (Fig. 5).

TCP segment format and simplifications

A TCP header and data block together are known as a TCP segment and has the following format [6, 7, 10]:

Source Port: 16 bits

The source port number. Parameter ignored in micro-implementation.

Destination Port: 16 bits

The destination port number. The microserver check the port field of every incoming TCP segment to see if it refers to the WWW service identified by a value of 80.

Sequence Number: 32 bits

The sequence number of the first data octet in this segment (except when SYN is present). If SYN is present the sequence number take the initial value and ACK (initial+1) as the first data octet (Fig. 5).

Acknowledgment Number: 32 bits

If the ACK control bit is set this field contains the value of the next sequence number the sender of the segment is expecting to receive. Once a connection is established this is always sent.

32-bit arithmetic is a problem on some microcontrollers because they support only 8-bit operations directly. Assume that the data stream is less than 64kB, and then is to perform only 16-bit operations on the 32-bit values.

Header Length: 4 bits

The number of 32 bit words in the TCP Header. This indicates where the data begins. The TCP header is an integral number of 32 bits long.

Reserved: 6 bits

Reserved for future use. Must be zero.

Control Bits: 6 bits

From left to right:

URG	(0x20):	Urgent Pointer field significant – not used in micro,
ACK	(0x10):	Acknowledgment field significant,
PSH	(0x08):	Push Function – not used in micro,
RST	(0x04):	Reset the connection,
SYN	(0x01):	Synchronize sequence numbers,
FIN	(0x01):	No more data from sender.

For more simplicity, the header length, reserved bits and flags are defined as two-byte values.

Window: 16 bits

The window size indicates the amount of Receive buffer space that is available for flow control. It can be set to a fixed size on transmit and ignored on receive.

Checksum: 16 bits

The TCP header must include a valid checksum value, which is computed across the whole TCP segment, plus a pseudoheader [6, 7, 10] containing not transmitted parts of the IP header. For simplicity, the checksum would be to compute on the fly as the header and data are sent out and then append to the data. While computing the checksum, the checksum field itself is replaced with zeros.

Urgent Pointer: 16 bits

This can be ignored because all data can be transmitted with equal priority.

Options: variable length

There is an addition variable-length header options field. Don't have to generate options, and discard any incoming options.

4. IP MICRO-IMPLEMENTATION

The Internet Protocol (IP) is used to transfer the TCP segments between hosts [11]. The IP is easiest to implement than the TCP protocol.

An IP header plus a TCP segment is known as a datagram and has the following format [6, 7, 11]:

Version: 4 bits

This section describes version IPv4.

Header Length: 4 bits

The default header size (measured in 32-bit words) is five.

Assume a value of 45h for the above parameters.

Type of Service: 8 bits

This field provides an indication of the abstract parameters of the quality of service desired. Prioritise datagrams by setting this field to zero, which is normal precedence.

Total Length: 16 bits

The length of the datagram, measured in octets, including Internet header and data.

Identification: 16 bits

An identifying value assigned by the sender to aid in assembling the fragments of a datagram. The value is incremented for each datagram sent.

Flags: 3 bits

The fragmentation procedure may be split a large datagram into smaller datagrams. The limitations of RAM on a microcontroller don't allow supporting fragmentation.

Fragment Offset: 13 bits

This field indicates where in the datagram this fragment belongs. The first has offset zero.

Time to Live: 8 bits

The constant value that is generated on transmit and ignored on receive.

Protocol: 8 bits

Identify which protocol (TCP or ICMP in micro) is used in the data area.

Checksum: 16 bit

A simple checksum of the IP header only.

Source and destination addresses: 2 x 32 bit

These are IP addresses expressed as 32-bit values. The microserver must be careful to use the incoming values in the outgoing datagrams.

5. SUMMARY

The above solutions to develop Web servers with less construction complexity and reduction of a power consumption are used in projects [8] and [9]. One implementation is based on 8-bit PIC16F877 microcontroller (Microchip) with 8k(14bit) FlashROM and 368 bytes of RAM. And the second implementation – does the microserver take advantage of AVR Atmega163 (Atmel), which has 16kB of FlashROM and 1024 bytes of RAM.

BIBLIOGRAPHY

- [1] Adamczyk A.: *Consumer Electronics – nowe wyzwanie dla Internetu*. Electroni Commerce – Gospodarka XXI Wieku, Wyd. Politechniki Gdańskiej, Gdańsk 2001, pp. 5-10.
- [2] Adamczyk A., Suchodolski W.: *Internetowa akwizycja sygnałów życiowych*. XII KK Biocybernetyka i inżynieria biomedyczna, 10-13 Sept. 2003, Gdańsk, vol. I, pp. 395-400.
- [3] Adamczyk A.: *Internetowa telediagnostyka urządzeń w obsłudze serwisowej*. Electronic Commerce – Teoria i Zastosowania, Wyd. Politechniki Gdańskiej, Gdańsk 2002, pp. 5-12.
- [4] Adamczyk A., Wierzbiński M.: *Internetowa telediagnostyka układów elektronicznych*. Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej, Seria: Technologie Informacyjne, 2003, no 1, pp. 245-252.
- [5] Parker T., Sportack M.: *TCP/IP Unleashed*. Sams Publishing, 2000.
- [6] Jeremy Bentham: *TCP/IP LEAN*. CMP Books, 2002.
- [7] Wright G. R., Stevens W. R.: *TCP/IP Illustrated, Vol. II*. Addison-Wesley Pub. Co. 1995.
- [8] Dec W. (promoted by Adamczyk A.): *Wielowojściowy rejestrator napięcia z mikroserwerem TCP/IP*. Graduate work 2003, Faculty of Electronic, Telecommunications and Informatics, Gdańsk University of Technology.
- [9] Jędrzyk R. (promoted by Adamczyk A.): *Mikroserwer TCP/IP realizujący sieciowe usługi kontrolno-pomiarowe zgodne z SCPI*. Graduate work 2003, Faculty of Electronic, Telecommunications and Informatics, Gdańsk University of Technology.
- [10] *Transmission Control Protocol*. RFC 793, Information Sciences Institute, University of Southern California, <ftp://ftp.rfc-editor.org/in-notes/std/std7.txt>
- [11] *Internet Protocol*. RFC 791, Information Sciences Institute, University of Southern California, <ftp://ftp.rfc-editor.org/in-notes/std/std7.txt>

IMPLEMENTACJA MIKROSERWERA**Streszczenie**

W artykule zaprezentowano zasady projektowania mikroserwerów z o ograniczonej złożoności obliczeniowej i niewielkim poborze energii. Zaproponowana redukcja funkcjonalności protokołów TCP/IP, pomimo niewielkiego zapotrzebowania na obliczenia, pozwala na realizację wydajnego połączenia klient-serwer w układach mikrokontrolerowych.

A. Czyżewski, A. Kaczmarek, P. Maziewski, P. Ody, P. Szczuko *
H. Krzyżniewski, L. Smoleński, M. Zientalski **

Katedra Systemów Multimedialnych, WETI, Politechnika Gdańska *
Katedra Systemów i Sieci Telekomunikacyjnych, WETI, Politechnika Gdańska **

BADANIE JAKOŚCI TRANSMISJI MOWY W SIECIACH IP

Streszczenie

Praca zawiera opis eksperymentu mającego na celu zbadanie relacji pomiędzy oceną subiektywną sygnału mowy a jakością transmisji tego sygnału w kanale telefonicznym VoIP. Wykorzystano symulację transmisji pakietowej sygnału w sieci IP. Wykonano serię testów odsłuchowych opartych na listach logatomowych i odpowiednio dobranych zdaniach. Do interpretacji wyników zastosowano analizę statystyczną.

1. WPROWADZENIE

W dobie rosnącego znaczenia sieci komputerowych, w tym globalnej sieci Internet, coraz większą popularność zyskują systemy umożliwiające transmisje dźwięków, a przede wszystkim mowy w tych sieciach. Niestety, dominujący obecnie protokół IP nie był projektowany pod kątem tego typu przekazów, co przekłada się na problemy z uzyskaniem odpowiedniej jakości transmitowanej w ten sposób mowy. Można oczekiwać, że usługa telefoniczna powinna być realizowana z jakością nie gorszą od uzyskiwanej w klasycznej sieci PSTN. Analogowy sygnał mowy jest przesyłany przez kanał utworzony w sieci IP (pomiędzy koderem i dekerem, analogicznie jak w PSTN) a zniekształcenia przesyłanego sygnału powodują pogorszenie jakości.

Istnieją liczne normy wskazujące na metody subiektywnej oceny jakości mowy [6, 8], brak natomiast metod obiektywnych. Opracowanie takich metod wydaje się być koniecznością ze względu na wymagania towarzyszące testom subiektywnym. Wymagania te w poważny sposób ograniczają możliwości szybkiego i nie wymagającego wysokich kosztów testowania m.in. kolejnych wersji kodeków mowy wykorzystywanych w VoIP a nawet poszczególnych parametrów kompresji.

VoIP – Voice over IP jest technologią służącą do przesyłania sygnałów fonicznych poprzez sieć pakietową. Jej głównym założeniem jest integracja ruchu telefonicznego z transmisją danych. Technologią VoIP zajmuje się wiele firm produkujących sprzęt i oprogramowanie dla telekomunikacji, a także większość operatorów telekomunikacyjnych, dlatego przewiduje się jej upowszechnienie. Obecnie technologia VoIP ma dwa podstawowe zastosowania. Pierwsze z nich to wykorzystanie jej do prowadzenia między-

narodowych lub międzymiastowych rozmów telefonicznych, zaś drugim zastosowaniem jest integracja sieci korporacyjnych. Połączenie sieci transmisji danych i transmisji mowy, powoduje obniżenie kosztów wdrożenia i utrzymania takiej sieci.

Na praktycznie uzyskiwaną jakość połączenia głosowego w technologii VoIP wpływ ma wiele czynników, takich jak: sposób kompresji mowy, powstawanie zjawiska echa, opóźnienie, jitter, zmiana kolejności i gubienie pakietów. W zgodnej opinii użytkowników zjawiska te znacząco pogarszają jakość transmisji, jednak brak jak dotychczas uniwersalnej, dobrze opracowanej metodyki oceny wpływu rodzaju kodowania i stopnia nasilenia niekorzystnych czynników na wynikową jakość transmisji.

W prezentowanym procesie oceny jakości transmisji sygnału mowy można wyodrębnić przygotowanie sygnałów wykorzystywanych do oceny, przesłanie sygnałów poprzez model kanału telefonicznego VoIP o zadanych parametrach i analizę sygnałów odebranych, zgodnie z przyjętą metodą oceny jakości sygnału mowy. Wyniki przedstawione w pracy uzyskano wykorzystując prosty model symulacyjny kanału telefonicznego VoIP, uwzględniający kodeki i przesyłanie pakietów w sieci IP, ze względu na łatwość realizacji i możliwość pełnej kontroli parametrów. Może on być zastąpiony przez system transmisji pracujący w rzeczywistej sieci. Model transmisji sygnałów VoIP przez sieć pakietową określa prawdopodobieństwo zdarzeń związanych z transmisją pakietów, bez uwzględniania struktury wewnętrznej sieci. Podstawowe parametry brane pod uwagę to prawdopodobieństwo utraty pakietu wewnątrz sieci, zmienność czasu przesyłania poszczególnych pakietów i prawdopodobieństwo powstania błędu binarnego w zawartości pakietu.

Podjęte badania mają na celu próbę zdefiniowania obiektywnych metod lub nowych norm. W niniejszej pracy przebadano cztery wybrane systemy kodowania sygnału mowy z wykorzystaniem symulowanego kanału telefonicznego. Parametrem testowanym było prawdopodobieństwo utraty pakietu. Taki wybór wiąże się z powszechnym uznaniem istotnego znaczenia tego parametru, jeśli chodzi o wpływ na zrozumiałość mowy [1, 2, 3, 5]. Zastosowano dwa rodzaje testów odsłuchowych: wyrazistości logatomowej i zrozumiałości zdaniowej. Wyniki obu testów należy traktować pilotażowo z uwagi na stosunkowo małą liczbę słuchaczy, z tym, że w oparciu o wyniki testu zrozumiałości zdaniowej można dokonać próby kategoryzacji.

2. PRZYGOTOWANIE NAGRAŃ

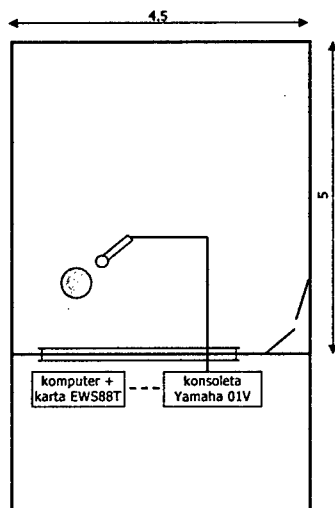
Przy wykonywaniu nagrań stosowano się do zaleceń zawartych w zaleceniach ITU [7]. Do nagrań wykorzystano pomieszczenie o objętości ok. 65 m³, którego czas pogłosu wynosi ok. 0.3 s. Poziom szumów nie przekraczał 30 dB. Pomiar wykonano z wykorzystaniem charakterystyki korekcyjnej typu A [4]. Materiał testowy rejestrowany był z wykorzystaniem mikrofonu Shure SM7, konsoli cyfrowej Yamaha 01V oraz komputera wyposażonego w kartę dźwiękową Terratec EWS88T. Pomiędzy konsolą a kartą dźwiękową dźwięk przysyłano cyfrowo z wykorzystaniem łącza SPDIF, co zapobiegało wprowadzeniu do dokonywanych nagrań ewentualnych zniekształceń i zakłóceń. Mikrofon umieszczony był w odległości ok. 15 centymetrów od ust osoby czytającej.

Nagrania wykonywane były z poziomem o ok. 20-25dB niższym od poziomem przesterowania, monofonicznie, z rozdzielczością 32 bitów i częstotliwością próbkowania 48 kHz. Po wstępnym montażu (wycięcie błędnie odczytanych kwestii) i normalizacji dokonano konwersji do parametrów zgodnych z płytą CD-Audio, tj. 16 bitów i 44,1 kHz.

Na materiał testowy składały się:

- logatomy – cztery listy stuwyrazowe
- zdania – w sumie 626 zdań o czasie odczytywania 2-3s

Materiał był odczytywany przez dwóch mężczyzn i dwie kobiety. Jest to liczba wystarczająca do oceny różnic wpływu transmisji mowy na głos męski i żeński.



Rys.1. Plan i wyposażenie pomieszczenia wykorzystanego do realizacji nagrań.
Wymiary pomieszczenia podano w metrach (4,5 x 5).

3. OPIS EKSPERYMENTU

W procesie symulowanej transmisji zastosowano cztery systemy kodowania sygnału mowy: ADPCM, ADPCM2, (przepływność dla obu 32 kbit/s, 4 bit/próbka), CELP (4,8 kbit/s) oraz GSM (13 kbit/s). Długości pakietów wynosiły odpowiednio: 152, 256, 12 i 65 bajtów. We wszystkich przypadkach pasmo zostało ograniczone do telefonicznego (częstotliwość próbkowania wynosiła 8 kHz). Symulacja dotyczyła zmiennego prawdopodobieństwa utraty pakietu w zakresie od 2% do 10% ze skokiem 2%.

W testach odsłuchowych uczestniczyło 5 osób (test logatomowy) i 4 osoby (test zdaniowy) nie znające zawczasu podawanych do odsłuchu zdań. Badanie zrozumiałości zdaniowej odbyło się zgodnie metodami ACR (Absolute Category Rating) [7], została jednak dokonana modyfikacja wykorzystująca sześciostopniową skalę oceny jakości transmisji – równoważniki opisowe kategorii:

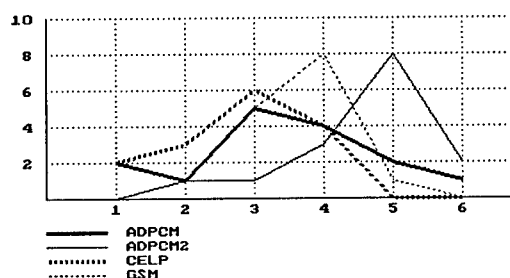
- | | |
|-------------------------|-------------------|
| 1 – niedostatecznie | 4 – dobrze |
| 2 – ledwie dostatecznie | 5 – bardzo dobrze |
| 3 – dostatecznie | 6 – doskonale |

Do przeprowadzenia testu logatomowego użyty został system kodowania CELP oraz 250 logatomów wybranych z 5 list logatomowych po 50 z każdej, natomiast test zrozumiałości zdaniowej został wykonany w oparciu o 60 zdań dla wszystkich czterech systemów kodowania i wszystkich zastosowanych wartości parametrów kanału. W tym drugim przypadku oprócz oceny kategorii jakości transmisji słuchacze notowali treść podanego do odsłuchu zdania, co pozwoliło im dokonać bardziej wnikliwej oceny treści.

3. WYNIKI

Wyniki testu logatomowego dla systemu kodowania CELP podano w tablicy 3.1.

parametru kanału w zależności od zastosowanego systemu kodowania (rys. 5 i tab. 3.7). Przy każdej tablicy podano wartość statystyki liczonej ze wzoru (3.1), liczbę stopni swobody i odpowiadającej jej wartości krytycznej na 5% poziomie istotności.



Rys.2. Rozkład ocen dla przykładowego słuchacza

Tablica 3.3

Rozkład ocen słuchacza nr 2

	1	2	3	4	5	6
ADPCM	2	1	5	4	2	1
ADPCM2	0	1	1	3	8	2
CELP	2	3	6	4	0	0
GSM	0	1	5	8	1	0

$\chi^2 = 30.333$ [liczba stopni swobody = 15], wartość krytyczna wynosi: 24.996 [5.0%]

Na podstawie otrzymanych wyników można wysnuć wniosek, że słuchacze kierowali się różnymi kryteriami (różne wykorzystania skali ocen), zaś niektórzy nie zauważyli różnic pomiędzy systemami kodowania (słuchacze 3 i 4 – różnice są nieistotne statystycznie na 5% poziomie istotności). Przyczyn takiego stanu rzeczy można upatrywać w zbyt małej liczbie nagrań (tylko 3 dla danego systemu i wybranej wartości parametru kanału) lub po prostu w zmęczeniu, czas trwania obu testów (łącznie z logatowym) wynosił ok. 2 godz.

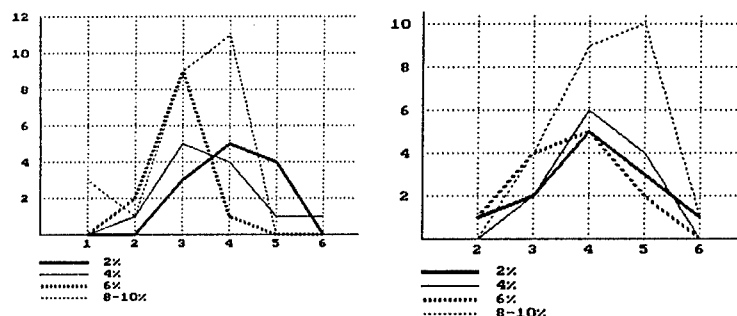
Zważywszy na powyższe wyniki można uznać te oceny za realizacje niezależnych zmiennych losowych i obliczyć ocenę łączną tak, jakby to było wykonane przez jednego eksperta, ale w oparciu zwielokrotnioną liczbę obiektów (z 60 otrzymuje się 240).

Tablica 3.4

Rozkład wszystkich ocen dla systemu kodowania ADPCM

	1	2	3	4	5	6
2%	0	0	3	5	4	0
4%	0	1	5	4	1	1
6%	0	2	9	1	0	0
>6%	3	1	9	11	0	0

$\chi^2 = 30.684$ [liczba stopni swobody = 15], wartość krytyczna wynosi: 24.996 [5.0%]

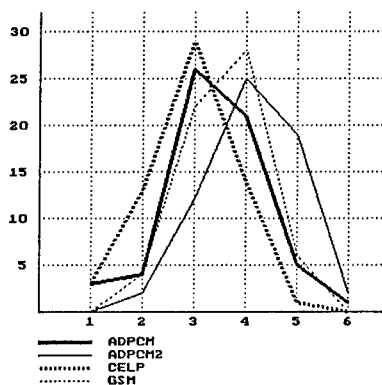


Rys.3. Rozkład wszystkich ocen dla systemu kodowania ADPCM.i ADPCM2

Tablica 3.5

Rozkład wszystkich ocen dla systemu kodowania ADPCM2

	2	3	4	5	6
2%	1	2	5	3	1
4%	0	2	6	4	0
6%	1	4	5	2	0
>6%	0	4	9	10	1

 $\chi^2 = 8.173$ [liczba stopni swobody = 12], wartość krytyczna wynosi: 21.026 [5.0%]


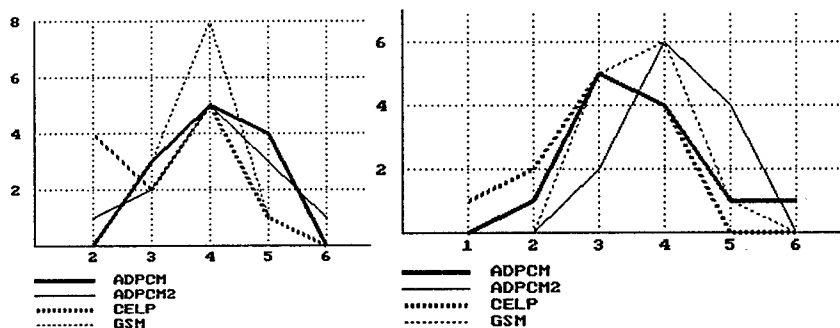
Rys.4. Rozkład wszystkich ocen dla poszczególnych systemów kodowania.

Tablica 3.6

Rozkład wszystkich ocen dla poszczególnych systemów kodowania

	1	2	3	4	5	6
ADPCM	3	4	26	21	5	1
ADPCM2	0	2	12	25	19	2
CELP	3	13	29	14	1	0
GSM	0	4	22	28	6	0

 $\chi^2 = 58.304$ [liczba stopni swobody = 15], wartość krytyczna wynosi: 24.996 [5.0%]



Rys.5 Rozkład ocen dla prawdopodobieństwa utraty pakietu = 2% i 4%

Tablica 3.7

Rozkład wszystkich ocen dla prawdopodobieństwa utraty pakietu = 2% i 4%.

prawdopodobieństwo	2%					4%					
	2	3	4	5	6	1	2	3	4	5	6
ADPCM	0	3	5	4	0	0	1	5	4	1	1
ADPCM2	1	2	5	3	1	0	0	2	6	4	0
CELP	4	2	5	1	0	1	2	5	4	0	0
GSM	0	3	8	1	0	0	0	5	6	1	0

 $\chi^2 = 16.174$ [liczba stopni swobody = 12], wartość krytyczna = 21.026 [5.0%] - dla 2%

 $\chi^2 = 18.055$ [liczba stopni swobody = 15], wartość krytyczna = 24.996 [5.0%] - dla 4%

4. DYSKUSJA WYNIKÓW I WNIOSKI

Na podstawie otrzymanych wyników nie jest jeszcze w pełni możliwe podejmowanie próby definiowania zakresów kategoryzacji dla poszczególnych systemów kodowania w zakresach zadawanych przez testowany model kanału. Wynika to z dużego rozrzutu ocen i w wielu przypadkach słabo zaznaczonych maksimów.

Można zauważyć, mimo tego, pewne tendencje dotyczące zwłaszcza odporności na utratę pakietów. Można uszeregować badane systemy kodowania w kolejności wynikającej z porównań statystycznych wyników testów subiektywnych z punktu widzenia prawdopodobieństwa utraty pakietu. Kolejność tę można ustalić na podstawie wartości obserwowanych statystyk zastosowanego testu niezależności. Mniejsza wartość tej statystyki od wartości krytycznej oznacza niezależność badanych cech na zadanym poziomie istotności (tutaj 5%). Jedną z tych cech jest typ zastosowanego systemu kodowania, drugą parametr kanału: prawdopodobieństwo utraty pakietu. Ze względu na różne liczby stopni swobody wartości krytyczne są różne i należy przy tych porównaniach obok wartości obserwowanych uwzględniać także te wartości krytyczne.

Ze względu na wartości obserwowanych statystyk można stwierdzić, że w testowanym zakresie parametru kanału nie stwierdzono różnic istotnych statystycznie dla systemu ADPCM2, jest więc on niewrażliwy na utratę pakietów od 2% do 10%, jednocześnie najbardziej wrażliwym jest system ADPCM. Z przeprowadzonych eksperymentów wynika, że najlepiej jest oceniany kodek ADPCM2, gorzej ADPCM. Pozostałe kodeki CELP i GSM są na tych wykresach ułożone w kolejności odmiennej niż na powyższej liście, co wskazuje na ogólną gorszą jakość kodeka CELP w stosunku do kodeka GSM, albo na słabą odporność na niski procent utraconych pakietów. Ponadto kodek CELP jest oceniany gorzej niż ADPCM. Wykres na rys. 10 obrazuje zbiorczą preferencję, potwierdzającą po-

wyższe obserwacje. Przyczyną niskiej oceny kodeka CELP jest zapewne największy stopień kompresji i co za tym idzie większa wrażliwość na utratę pakietu.

Tablica 3.9 stanowi prezentację możliwości kategoryzacji testowanego parametru kanału transmisyjnego. Obserwowane maksima na wykresach powinny być podstawą dla określenia zakresów parametru dla zdefiniowanych kategorii subiektywnych. Pomimo dużego rozrzutu ocen, a w niektórych sytuacjach szerokich maksimów na wykresach preferencyjnych, dają się zauważyć pewne tendencje w oparciu o analizę statystyczną.

Tablica 3.9

Wyniki testów niezależności ocen od parametru kanału dla testowanych systemów kodowania

Kodowanie	χ^2	wartość krytyczna
ADPCM2	8.173	21.026
CELP	21.266	21.026
GSM	19.610	16.919
ADPCM	30.684	24.996

Obserwuje się, bowiem, wzrost wartości statystyki określającej niezależność pomiędzy ocenami poszczególnych systemów ze wzrostem liczby utraconych pakietów w kanale. Oznacza to, że różnice pomiędzy systemami kodowania uwidaczniają się lepiej dla większej liczby tego typu zakłóceń transmisji (większego prawdopodobieństwa utraty pakietu). Dodatkowo potwierdzają się wyżej podane wnioski, odnośnie ogólnej oceny testowanych systemów w warunkach zamodelowanych w opisywanym eksperymencie.

BIBLIOGRAFIA

- [1] Brachmański S.: *Ocena jakości transmisji mowy w sieciach IP metodą ACR*, Prace 50 Otwartego Seminarium z Akustyki, str.480-483, Wrocław, 2003.
- [2] Brachmański S.: *VoIP – ocena jakości transmisji mowy metodą ACR i DCR*, PtiWT 8-9/2003, s.424-7.
- [3] Bachmański S., Kula S.: *Badanie jakości mowy w połączeniach głosowych. Stara usługa – nowe problemy*, PtiWT 8-9/2003, s.418-23.
- [4] Davis G., Jones R.: *Sound Reinforcement Handbook*. Yamaha Corporation of America and Gary Davis & Associates, June 1989.
- [5] Gozdecki J., Jajszczyk A., Stankiewicz R.: *QoS Terminology in IP Networks*. IEEE Comm. Mag. 41(3), 2003, 153-159.
- [6] ITU-T P.861, *Objective Quality Measurement of Telephone-Band (300-3400) Speech Codecs*, Geneva, Switzerland, August 1996.
- [7] ITU-T Recommendation P.800 *Method for subjective determination of transmission quality*, 1996
- [8] ITU-TS Recommendation P.59 – *Artificial Conversational Speech*. Geneva, March 1993.

INVESTIGATIONS OF QUALITY OF SPEECH TRANSMITTED OVER IP NETWORKS

Summary

The paper contains a description of an experiment investigating the relation between a subjective speech rating and an objective quality of a telephone channel VoIP. An Internet packet transmission conditions were simulated. Listening tests were carried out based on non-sense syllable lists and a set of common speech sentences. The statistical processing of results allowed to draw-out and to present some general conclusions concerning the influence of transmission channel parameters on the subjective assessment of speech quality.

Piotr Gajos, Sławomir Kwiatkowski

DGT Spółka z o.o.

URZĄDZENIA FIRMY DGT W TECHNOLOGII VOIP I ICH INSTALACJA W PILOTAŻOWEJ SIECI

Streszczenie

W publikacji przedstawiono najnowsze urządzenia produkcji DGT zrealizowane w technologii VoIP. Technologia przesyłania głosu przez Internet jest dość nowym rozwiązaniem znajdującym coraz częstsze zastosowanie komercyjne. Rozwiązania e-biznesowe wzbogacone o usługę przesyłania głosu za pośrednictwem protokołu IP z jednoczesną możliwością współpracy z centralkami biurowymi oraz publiczną siecią telekomunikacyjną cieszą się obecnie dużym zainteresowaniem wśród odbiorców systemów nowej generacji. Instalacja pilotażowa systemu VoIP u jednego z operatorów telewizji kablowej pozwoliła zdobyć wiele istotnych doświadczeń. Opracowanie koncepcji, realizacja i wdrożenie tego systemu stały się dla DGT cennym źródłem informacji, którą przedstawiamy w niniejszym materiale. Z uwagi na z góry ograniczoną objętość publikacji zagadnienia opisujące realizację instalacji pilotażowej są ogólnej natury.

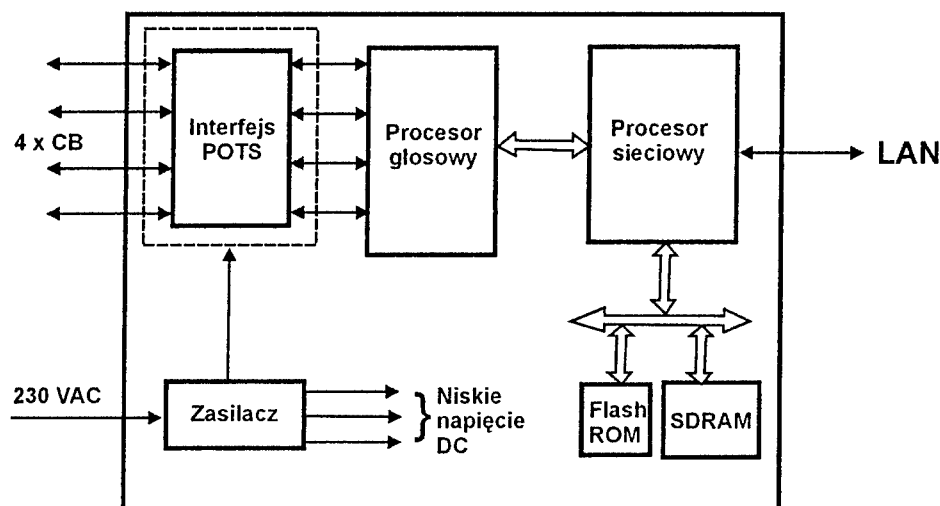
1. WSTĘP

Wielu odbiorców poszukuje platform komunikacyjnych oferujących dostarczanie zaawansowanych rozwiązań e-biznesowych wzbogaconych o usługę przesyłania głosu za pośrednictwem protokołu IP z jednoczesną możliwością współpracy z centralkami biurowymi oraz publiczną siecią telekomunikacyjną. Nowoczesne rozwiązania teleinformatyczne oferują również integrację z systemami CRM (*Customer Relationships Management*). Wszystko to sprzyja obniżaniu kosztów połączeń międzyoddziałowych z efektywnym wykorzystaniem zasobów sieci WAN.

Użytkownicy posiadający własne sieci teleinformatyczne oraz łączy transmisji danych zestawione pomiędzy różnymi odbiorcami mogą uzyskiwać znaczne korzyści dzięki zastosowaniu technologii VoIP. Niższe koszty urządzeń, utrzymania sieci oraz personelu, a przede wszystkim opłat za rozmowy telefoniczne to główne zalety telefonii w sieciach IP. Analiza rynku oraz doświadczenia użytkowników wskazują na potrzebę stosowania takich urządzeń, aby osiągnąć bardzo krótki czas zwrotu inwestycji. Ten czas to zazwyczaj okres kilkunastu miesięcy, co czyni inwestycje tego typu ciekawą propozycją poszerzenia oferty np. dla operatorów telewizji kablowych (TVK), dostawców internetu (ISP) oraz szerokich zastosowań u klientów biznesowych. Taka sieć VoIP może być wyposażona w urządzenia DGT takie jak ATK oraz SKB, które są szerzej opisane w dalszej części artykułu.

2. ABONENCKI TERMINAL KABLOWY DGT 7410

Abonencki Terminal Kablowy DGT 7410 (ATK) jest urządzeniem dedykowanym dla klientów (operatorów), którzy na bazie lokalnej sieci pragną dodatkowo świadczyć usługi przesyłania głosu. Do strony abonenta można podłączyć 4 standardowe analogowe aparaty telefoniczne (również faks) w celu wykonywania tradycyjnych połączeń telefonicznych oraz przesyłania dokumentów faksem. Konstrukcja Abonenckiego Terminala Kablowego (Rys.1) i użyte w nim podzespoły zapewniają maksimum niezawodności dla tego rozwiązania. Czas jego bezawaryjnej pracy jest obecnie szacowany na okres 15 lat.



Rys.1. Schemat blokowo-funkcjonalny Abonenckiego Terminala Kablowego DGT 7410

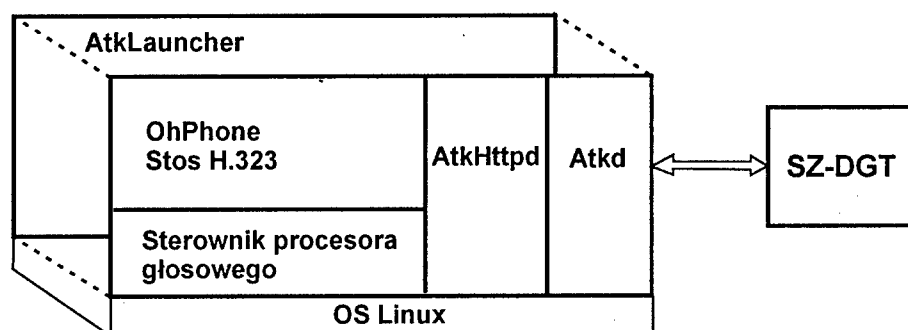
Moduł interfejsu POTS w ATK jest odpowiedzialny za przenoszenie sygnalizacji oraz mowy pomiędzy łączem obsługującym telefony analogowe a łączem cyfrowym procesora głosowego. Procesor głosowy jest odpowiedzialny za próbkowanie oraz konwersję pakietów głosowych. Procesor sieciowy dokonuje pakietyzacji strumienia danych głosowych i przesyłania ich zgodnie z protokołem H.323 w sieci LAN. Do takich operacji jest wykorzystywana pamięć operacyjna SDRAM. Z pamięci Flash przy każdorazowym uruchomieniu urządzenia jest ładowane oprogramowanie niezbędne do funkcjonalności urządzenia.

Po dołączeniu do ATK standardowego telefonu analogowego możliwa jest realizacja połączeń głosowych poprzez sieci IP, jak również pomiędzy abonentami sieci IP i abonentami sieci PSTN. ATK współpracuje z Serwerem Komutacji Brzegowej (SKB), który realizuje dostęp do sieci telekomunikacyjnych PSTN i innych sieci abonenckich. Można powiedzieć, że urządzenia te pełnią funkcję bram umożliwiających zestawianie połączeń pomiędzy dwoma do niedawna odrębnymi sieciami: siecią telekomunikacyjną (PSTN) a siecią przesyłania danych opartą o protokół IP. Komunikacja pomiędzy ATK a SKB odbywa się poprzez sieć Ethernet.

Urządzenia ATK o przedstawionej wyżej budowie umożliwiają realizację następujących funkcji:

- niezależny i równoczesny dostęp do usługi POTS dla 4 abonentów posiadających aparaty telefoniczne analogowe lub faksy,
- zestawianie kanałów transmisji głosu z wykorzystaniem konwersji w standardzie G.711, G.729 A/B lub G.723.1 (możliwe dla wszystkich portów jednocześnie),
- przesyłanie głosu poprzez Ethernet w standardzie H.323 w oparciu o protokół IP,
- optyczną sygnalizację stanu pracy urządzenia (diody LED),
- możliwość konfiguracji urządzenia poprzez sieć TCP/IP (strona WWW), przez port RS (tryb tekstowy) lub z poziomu Systemu Zarządzania DGT (SZ DGT),
- eliminację echa do 64 ms zgodnie ze standardem G.168.

Na rysunku 2 została przedstawiona struktura oprogramowania ATK osadzona na platformie OS Linux.



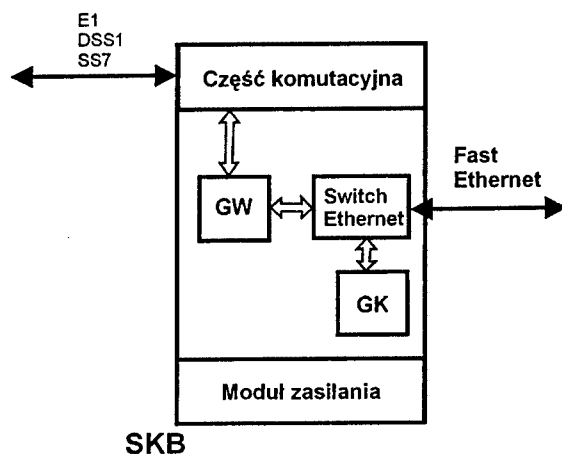
Rys. 2. Struktura modułowa oprogramowania ATK

AtkLauncher jest modulem uruchamiającym oraz nadzorującym pozostałe aplikacje takie jak:

- OhPhone – moduł odpowiedzialny za obsługę portu abonenckiego oraz konwersję na pakiety H.323,
- AtkHttpd – moduł odpowiedzialny za konfigurację urządzenia,
- Atkd – klient komunikujący się z SZ DGT,
- Sterownik procesora głosowego - sterownik obsługujący procesor głosowy.

3. SERWER KOMUTACJI BRZEGOWEJ - SKB

Serwer Komutacji Brzegowej – SKB jest instalowany w punkcie styku sieci IP i PSTN, co pozwala na realizację usług dotychczas dostępnych tylko w obrębie jednej z nich. Na przykład w sieci transmisji danych opartej na protokole IP można bez zmiany infrastruktury korzystać z wielu dodatkowych usług, takich jak telefonia tradycyjna w sieci PSTN, VoIP, dostęp do Internetu. SKB jest urządzeniem, który od strony sieci IP może obsłużyć i zarządzać grupą nawet kilkudziesięciu tysięcy terminali. Od strony sieci PSTN jest wyposażony w bogaty zestaw interfejsów i sygnalizacji tak, aby można go było połączyć do każdej centrali czy systemu telekomunikacyjnego. Przykładowe elementy składowe serwera SKB i sposób połączenia z sieciami oraz sama struktura jego modułów została pokazana na rys. 3.

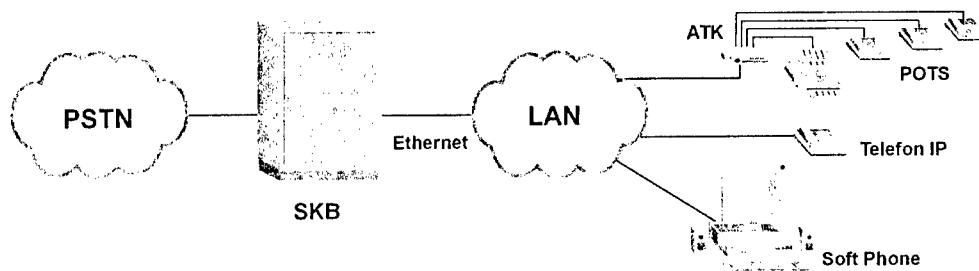


Rys. 3. Struktura i sposób połączenia modułów SKB

W skład SKB wchodzi następujące elementy:

- **Część komutacyjna** realizująca zestawianie połączeń na styku z siecią PSTN. Połączenia te mogą być realizowane do sieci PSTN jak również do sieci korporacyjnych lub samych central PABX
- **GW** – moduł gateway’a jest odpowiedzialny za konwersję pakietów sygnalizacyjnych oraz głosowych pomiędzy siecią IP a częścią komutacyjną
- **GK** – gatekeeper realizuje funkcję zestawiania połączeń pomiędzy użytkownikami sieci IP jak i również pomiędzy użytkownikami sieci IP i PSTN.
- **Switch Ethernet** jest urządzeniem wbudowanym w SKB
- **Moduł zasilania** doprowadza zasilanie do wszystkich modułów i elementów

Dzięki wszechstronności SKB możliwa jest obsługa różnych terminali umożliwiających przesyłanie głosu poprzez sieci IP. Do tych urządzeń, poza ATK, można zaliczyć osobiste komputery z oprogramowaniem telefonicznym (Soft Phone), lub telefony IP (IP Phones). Urządzenia te mogą być włączane bezpośrednio do sieci informatycznej. SKB może obsługiwać też, bezpośrednio do nich dołączane, standardowe aparaty telefoniczne. Przykładowa taka sieć została pokazana na rys.4.

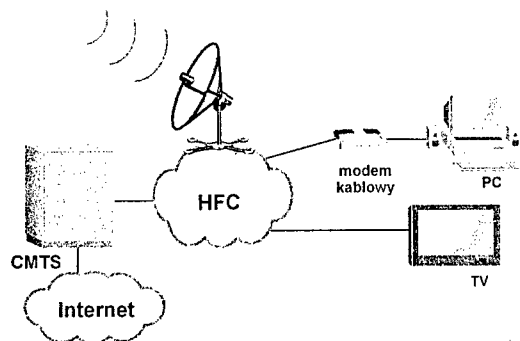


Rys.4. Schemat sieci VoIP w oparciu o rozwiązania DGT

SKB pozwala na łatwą, szybką i niedrogą implementację nowych usług na bazie istniejącej infrastruktury sieci transmisji danych. Zaimplementowanie sygnalizacji SS7 umożliwia połączenie serwera z wieloma operatorami jednocześnie. Dla wszystkich typów sygnalizacji można ustawiać parametry oraz określać zakres informacji przekazywanych do central lub innych urządzeń współpracujących z SKB. W serwerze można tworzyć grupy abonentów wirtualnych i dla każdego z nich może być przydzielony indywidualny licznik taryfikacyjny. SKB może być zarządzany lokalnie lub zdalnie z wykorzystaniem przeglądarki HTML, współpracującej z klasami Javy. Zaletą jest polski interfejs użytkownika. Serwer można dowolnie konfigurować, co pozwala na elastyczne dobranie jego parametrów do potrzeb klienta. SKB może służyć jako brama (gateway) między sieciami IP i PSTN realizując połączenia międzynarodowe lub międzymiastowe oraz do innych operatorów. Specjalne mechanizmy zarządzania ruchem pozwalają na optymalne gospodarowanie łączami dzierżawionymi (IP). SKB może też obsługiwać np. korporacyjne centrale PABX. W dużych i rozległych sieciach można tworzyć grupy SKB.

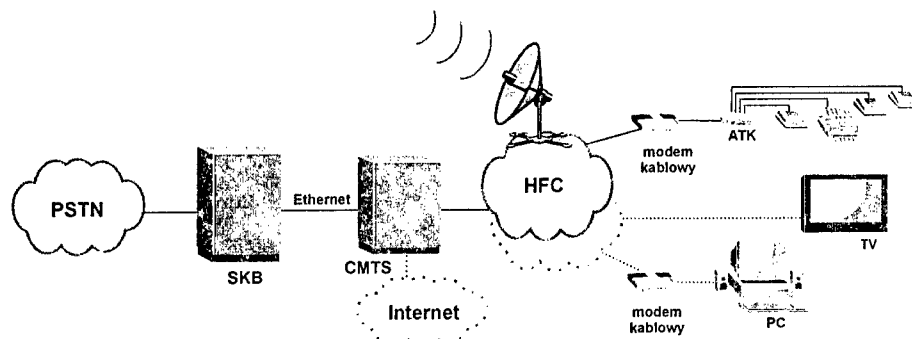
4. DOŚWIADCZENIA Z REALIZACJI INSTALACJI PILOTAŻOWEJ U OPERATORA TELEWIZJI KABLOWEJ

Od samego początku prace dotyczące firmowych rozwiązań VoIP toczyły się w porozumieniu i przy czynnej współpracy z potencjalnymi nabywcami. Ukoronowaniem tego etapu było uruchomienie instalacji pilotażowej a następnie instalacji komercyjnej w sieci Telewizji Kablowej firmy Gawex Media w Szczecinku. Proces uruchamiania trwał kilka miesięcy i był źródłem wielu cennych spostrzeżeń i doświadczeń. Podstawowym założeniem prac było maksymalne wykorzystanie już istniejącej infrastruktury i urządzeń pracujących w sieci Telewizji Kablowej, (TVK) które, oprócz dostarczania sygnału telewizyjnego pozwalały operatorowi TVK oferować swoim klientom możliwość dostępu do internetu za pośrednictwem tego samego okablowania. W węzle operatora TVK firma DGT uruchomiła Serwer Komutacji Brzegowej (SKB), zaś w wybranych punktach zainstalowano ATK z interfejsami pozwalającymi na podłączenie standardowych aparatów telefonicznych. Już pierwsze próby urządzenia wykazały, że ATK pozwala na uzyskanie wysokiej jakości dźwięku, również w przypadku zastosowania kompresji mowy. Ruch generowany w sieci IP okazał się relatywnie mały. Sieć operatora telewizji kablowej przed instalacją technologii VoIP została pokazana na rys. 5.



Rys. 5 Schemat sieci operatora telewizji kablowej

Ze względu na specyfikę sieci TVK z kanałem zwrotnym zdecydowano się oddzielić logicznie sieć VoIP od sieci dostarczającej Internet (Rys. 6). Dzięki temu można było się



Rys. 6. Schemat sieci operatora telewizji kablowej po instalacji rozwiązań technologii VoIP

skupić na pokonywaniu problemów pojawiających się w wyższych warstwach, traktując warstwę transportową jako na tym etapie nieistotną. Jak się okazało było to rozumowanie błędne. W związku z tym, że ATK w wersji DGT 7410 wyposażone jest w jeden port Ethernet, w jednym z węzłów dla podłączenia większej ilości urządzeń zainstalowano hub. Ponieważ był on wyjątkowo niskiej jakości to stał się źródłem trudnych do odtworzenia w laboratorium DGT problemów. Zwróciło to uwagę na konieczność stworzenia narzędzi diagnostycznych dla sieci, dedykowanych specjalnie do diagnozowania tego rodzaju zagadnień. Uproszczoną analizę określonego segmentu sieci może wykonać wpięty do niej komputer PC z odpowiednim oprogramowaniem. Okazało się bowiem, że nowoczesność technologii VoIP powoduje, że nawet wyrafinowane przyrządy pomiarowe, nie bardzo potrafią analizować sieci VoIP, chociaż taką funkcjonalność oferują.

Pierwsze egzemplarze urządzeń ATK 7410 były wyposażone w możliwość programowania oraz konfiguracji wyłącznie lokalnie przez port RS. Podczas rozbudowy instalacji pilotażowej i intensywnego rozwoju oprogramowania ATK okazało się, że nie jest to rozwiązanie zbyt wygodne. W związku z tym przyspieszono prace nad integracją ATK z Systemem Zarządzania DGT (SZ DGT). W efekcie, w krótkim czasie powstało narzędzie pozwalające nie tylko diagnozować stan podłączonych urządzeń, ale także zdalnie aktualizować oprogramowanie nawet dla dużej grupy zainstalowanych urządzeń, bez blokowania sieci i powodowania utrudnień komunikacyjnych dla abonentów. Specyficzną cechą SZ DGT jest możliwość zdefiniowania maksymalnej przepustowości sieci wykorzystywanej dla potrzeb aktualizacji oprogramowania. Zostało to podyktowane faktem, że nie wszystkie urządzenia w sieci pilotażowej spełniały wymagania QoS, co mogło powodować zablokowanie sieci podczas wykonywania takiej operacji. Oczywiście w sieci z zapewnionym QoS wydaje się, że problem ten nie wystąpi.

Kolejnym zagadnieniem była ocena i decyzja, co do konieczności dublowania urządzeń sterujących ruchem w sieci IP. Doświadczenia DGT z ostatnich lat pokazują, że, dzięki wysokiemu poziomowi niezawodności zastosowanych sterowników mikroprocesorowych nie wymagają one gorącej rezerwy do zachowania odpowiednich parametrów pracy. Istotnym zagadnieniem jest jednak zdalna aktualizacja wersji oprogramowania. O ile w części komutacyjnej DGT dopracowało się własnych, autorskich rozwiązań, pozwalających na aktualizację kodu podczas pracy urządzeń, to w przypadku rozwiązań VoIP jest jeszcze wiele do zrobienia. Ze względu na specyfikę technologii pakietowej, zagadnienia te

są w DGT realizowane łącznie w ramach jednego tematu, czyli rozpraszanie i równoważenie obciążenia sterowania w obrębie SKB.

Uruchomiona przez DGT instalacja pilotowa została włączona do sieci publicznej. Styk z PSTN, poprzez E1 z sygnalizacją DSS1 został zapewniony przez firmę Tel-Energo S.A. – jednego z operatorów. W trakcie uruchamiania okazało się, że poważnym problemem jest pojawiające się na tym styku echo będące niekorzystnym zjawiskiem już na etapie oceny subiektywnej. Docierające echo miało na tyle niekorzystny charakter (bardzo duże opóźnienie), iż nie udało się go wyeliminować za pomocą tradycyjnych urządzeń. W efekcie moduł bramy VoIP zastosowany w SKB nie był w stanie go wyeliminować. W trakcie rozwiązywania problemu przetestowano gateway'e różnych producentów, które pomimo deklarowanych parametrów nie były w stanie poradzić sobie z eliminacją tego problemu. Dopiero zastosowanie specjalistycznego eliminatora echa jednej z renomowanych firm pomogło w rozwiązaniu problemu. Podczas wnikliwej analizy okazało się, iż echo jest generowane przez część central funkcjonujących w sieci PSTN. Odbicie nadawanego dźwięku dla użytkowników sieci PSTN jest na tyle małe, iż nie są oni w stanie stwierdzić czy dźwięk, który słyszą jest generowany przez nich samych czy pochodzi od elementu, w którym powstaje odbicie. Niestety inaczej jest w sieci IP, ponieważ jak wiadomo, w tego typu sieciach występuje znacząco większe opóźnienie niż w sieci PSTN. Opóźnienie to stwarza poważny problem echa, który należy eliminować w celu świadczenia wysokiej jakości usług telefonicznych VoIP.

Kolejnym ciekawym zagadnieniem okazało się samo wybieranie numeru abonenta żadanego. ATK przyjmując kolejne cyfry za pomocą DTMF nie potrafiło samo stwierdzić ile powinien on mieć cyfr. W efekcie przesyłało cały numer do SKB dopiero po rozpoznaniu 6 sekund ciszy. Dla eliminacji tej uciążliwości wprowadzono dodatkowe rozwiązanie polegające na transmitowaniu do SKB informacji o numerze w układzie cyfra po cyfrze. Dzięki temu SKB mógł na podstawie odpowiednich tablic sam określić kiedy numer jest kompletny, a następnie natychmiast rozpocząć zestawianie połączenia. Jak się jednak okazało – niektóre gateway'e (wykorzystane do współpracy z siecią PSTN) nie oferowały tego trybu pracy.

Ostatecznie duży wysiłek włożony w realizację instalacji pilotażowej zaowocował przekształceniem jej w instalację komercyjną.

Podsumowując należy stwierdzić, że technologia VoIP dojrzeła do powszechnych zastosowań biznesowych i publicznych. Prawidłowe działanie zapewnione jest często wyłącznie w ramach własnych urządzeń. Dostawcy starają się jednak te problemy eliminować, co w niedalekiej przyszłości spowoduje pełną zgodność urządzeń w ramach obowiązujących standardów.

5. DALSZE PRACE NAD TECHNOLOGIĄ VOIP

W trakcie prac przy ATK pojawiło się kilka zagadnień, które firma DGT uznała za interesujące do bliższego rozpoznania i być może implementacji. Aktualnie w DGT w dziedzinie technologii VoIP są realizowane następujące tematy:

- Klastrowa struktura serwerów zgodnie z zaleceniem ITUT H.323,
- Implementacja protokołu SIP na platformie sprzętowej ATK DGT 7410,
- Implementacja protokołu MEGACO/H.248 na platformie sprzętowej ATK DGT 7410.

Bardzo duży nacisk przy formułowaniu założeń poszczególnych tematów położono na przygotowanie pracy według standardów i norm europejskich.

Dzięki bliskiej współpracy z Politechniką Gdańską udało się tymi zagadnieniami zainteresować kadrę WETI. Zaowocowało to listą propozycji tematów Prac Dyplomowych (o różnym poziomie złożoności) przeznaczonych do rozwiązania przez studentów.

Wszystkie te tematy zostały doprecyzowane przy współpracy konstruktorów DGT oraz promotorów z Wydziału ETI Politechniki Gdańskiej. Dzięki czemu dla tej grupy studentów może okazać się to pierwszą pracą, która zostanie zastosowana w przemyśle.

Wyposażono laboratoria Wydziału Elektroniki Telekomunikacji i Informatyki Politechniki Gdańskiej w niezbędną ilość Abonenckich Terminali Kablowych DGT 7410 dla realizacji zajęć dydaktycznych. Jednocześnie też przedstawiono szczegółowo dotychczasowe rozwiązania firmowe i zaprezentowano wizję dalszego rozwoju prac nad zastosowaniem technologii VoIP. Współpraca w tym zakresie jest bardzo obiecująca, zarówno dzięki dużemu zaangażowaniu kadry Wydziału ETI Politechniki Gdańskiej w proces dydaktyczny, jak też poprzez bardzo obiecujące efekty dotychczasowych prac studentów nad powierzonymi zadaniami.

BIBLIOGRAFIA

- [1] RFC2848 *PSTN and Internet Internetworking*.
- [2] RFC3136 *Services in the PSTN in Requesting Internet Services*.
- [3] ITU-T H.323 (2003-07) *Packet-based multimedia communications systems*.
- [4] RFC3435 Version 1.0 (2003-01) *Media Gateway Control Protocol (MGCP)*.
- [5] A. Jajszczyk: *Integracja usług w sieciach IP*. IV Ogólnopolska Konferencja „Warsztaty ISDN – Wrocław 2001”.
- [6] K. Lysek: *Technika IP w sieciach dostępowych*. V Ogólnopolska Konferencja „Warsztaty ISDN – Wrocław 2002”.
- [7] DGT Sp. z o.o.: *Serwer Komutacji Brzegowej, Abonencki Terminal Kablowy*. Suplement promocyjny. Produkty sieciowe high-tech (18) - kwiecień 2003.
- [8] M. Brejza: *Rodzime rozwiązania w natarciu*. Infotel – grudzień nr12/2003.
- [9] W. Urbanek: *Przebudzenie trójmiejskiej telekomunikacji*. Świat Telekomunikacji - styczeń 2004.
- [10] P. Szczepaniak: *Kablofonia IP po Polsce*. Networld nr 3/2004.

DGT PRODUCT-DEVICES SERVICING THE VOIP TECHNOLOGY AND THEIR INSTALLATION IN THE PILOT NETWORK

Summary

This publication presents the latest product-devices, servicing VoIP technology made by DGT. The technology of voice transmitting via Internet is relatively new solution more and more applicable on the market. E-business solutions, enriched with the service of voice transmission over the IP protocol, featured by simultaneous ability to work with office exchanges and PABX network, enjoy success among the clients of telecommunication systems of new generation. The pilotage installation of VoIP system in the base of one of cable-television operators has enabled us to gain lots of substantial experience. Working-out of concept, project accomplishment and implementation of such a system has become the valuable source of information, which is provided, in the following publication by DGT. Taking into account limited volume of the material in advance, the issues describing pilotage installation are of general nature.

Tomasz Gierszewski

Katedra Systemów Informacyjnych, Politechnika Gdańska

ODPORNOŚĆ NA AWARIE SIECI BEZSKALOWYCH

Streszczenie

W referacie dokonano porównania odporności na awarie sieci losowych oraz bezskalowych. Zaprezentowano dwa modele awarii – awarie części węzłów sieci wybranych losowo bądź rozmyślnie. Skomentowano rozbieżność w odporności na awarie sieci losowych oraz bezskalowych. Badania przeprowadzono dla różnych kryteriów funkcjonalności sieci. Zaprezentowano też kryterium pozwalające porównywać różne sieci między sobą.

1. WSTĘP

Internet miał być, zgodnie z intencją jego twórców, siecią odporną na awarie. Projekt badawczy zakładał konstrukcję sieci komputerowej odpornej na uszkodzenia znaczącej liczby węzłów. Przyjmuje się, że takie kryterium spełniają tak zwane sieci losowe. Właściwości sieci losowych zostały sformułowane przez parę węgierskich uczonych – Erdősa i Rényi'ego na początku lat sześćdziesiątych [1]. Topologia dzisiejszej sieci Internet¹ ma, jak się okazuje, charakter inny niż typowa topologia sieci losowych. Połączenia między routerami należącymi do różnych systemów autonomicznych są przykładem sieci określanej obecnie mianem bezskalowej (ang. *scale-free*). Właściwości tych sieci są szczególnie interesujące z powodu innych zachowań w przypadku np. rozmyślnego uszkodzenia.

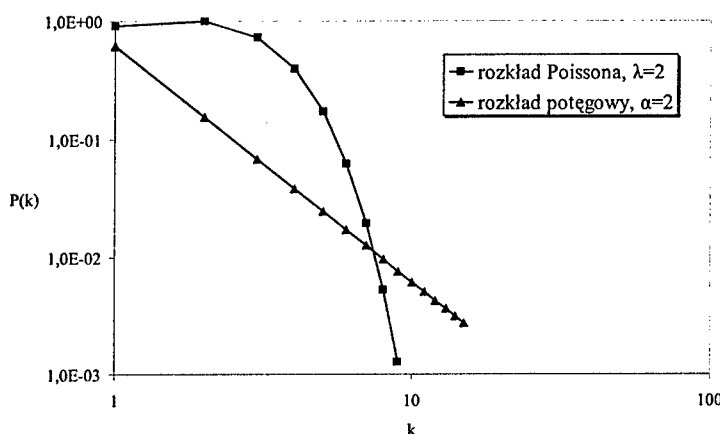
2. DEFINICJA SIECI BEZSKALOWYCH

Sieci losowe są często nazywane sieciami eksponencjalnymi. Nazwa ta bierze się stąd, że prawdopodobieństwo posiadania przez węzeł połączenia z k sąsiadami osiąga maksimum dla wartości średniej k (dla wszystkich węzłów sieci) oraz maleje wykładniczo dla dużych k , przy założeniu że liczba wszystkich połączeń w sieci jest liniowo zależna od liczby węzłów. Liczba połączeń z sąsiadami nazywana jest stopniem węzła i oznaczana przez k . Rozkład liczby węzłów w zależności od stopnia ma charakter rozkładu Poissona

¹ połączenia za pomocą routerów odrębnych systemów autonomicznych

(rys. 1). Powoduje to niezwykle rzadkie (praktycznie brak) występowanie węzłów o liczbie połączeń znacząco różnej od średniej.

W odróżnieniu od sieci losowych, stopień węzłów w sieciach bezskalowych charakteryzuje rozkład potęgowy. Oznacza to, że prawdopodobieństwo posiadania przez węzeł k sąsiadów jest proporcjonalne do $k^{-\gamma}$. W sieciach z takim rozkładem charakterystyczne jest występowanie węzłów o dużej w porównaniu ze średnią liczbie krawędzi. Węzły te nazywane są centrami. Nazwa sieci bezskalowe pochodzi od rozbieżności między dominującą ilościowo w sieci wartością stopnia, a wartościami charakterystycznymi dla centrów. Obecność centrów powoduje, że własności sieci bezskalowych różnią się znacznie od własności sieci losowych. Między innymi różna jest odporność na awarie tych dwóch rodzajów sieci.



Rys. 1. Porównanie rozkładów Poissona i potęgowego

Bezskalowy charakter różnych sieci zaobserwowano w wielu dziedzinach nauki [2], [3]. Z punktu widzenia teleinformatyki najbardziej interesujące są własności sieci informatycznych. Korzystając z wyników badań niezależnych zespołów [4], [5], [6], [7] można stwierdzić, że topologia połączeń między routerami najwyższego poziomu (poziomu systemów autonomicznych) ma charakter sieci bezskalowej.

3. RODZAJE USZKODZEŃ

W niniejszym referacie badane są własności sieci bezskalowych w porównaniu do sieci losowych. Typową rozbieżnością jest inne zachowanie obu typów sieci w przypadku awarii węzłów. Różnice uwypuklają się, jeśli uszkodzenia podzielone zostaną na przypadkowe awarie i celowe ataki.

W kontekście sieci losowych w zasadzie trudno mówić o różnicach między rozmyślnym atakiem a przypadkową awarią. Wynika to z braku występowania centrów, czyli węzłów, które mają połączenia ze stosunkowo dużą liczbą sąsiadów. Oczywiście w konkretnych przykładach sieci losowych można wyróżnić węzły o bardziej i mniej krytycznym znaczeniu. Jednak statystyczne badania sieci losowych nie ujawniają szczególnych prawidłowości.

Inaczej przedstawia się sytuacja w sieciach bezskalowych. Centra mają w tych sieciach krytyczne znaczenie. Eliminacja centrów może doprowadzić sieć w dużo krótszym czasie do niefunkcjonalności. Dlatego też należy zdefiniować co będzie oznaczało awarię, a co atak.

3.1. Uszkodzenie losowe – awarie

Losowe uszkodzenia mogą być pojmowane jako przypadkowe awarie sieci. Sieci bezskalowe okazują się w tym wypadku nieco odporniejsze niż sieci losowe. Wynik taki można łatwo wytłumaczyć – porównywane sieci mają tę samą liczbę węzłów i krawędzi. Skoro w sieciach bezskalowych istnieją centra, skupiają one pewien stosunkowo duży odsetek krawędzi. Tym samym pozostałe węzły mają ich przeciętnie nieco mniej.

Symulacja uszkodzenia była realizowana przez losowanie z rozkładem równomiernym kolejnych węzłów z zestawu aktualnie dostępnych (jeszcze nieusuniętych).

3.2. Uszkodzenie celowe – ataki

Celowe uszkodzenia miały symulować ataki obliczone na największe zniszczenia, jakie w rozmyślny sposób można dokonać w sieci przez eliminację węzłów. Eliminowane były więc centra. Ponieważ na skutek stopniowego usuwania wierzchołków z sieci zmieniają się też stopnie pozostałych węzłów, należało to uwzględnić w symulacji. Po każdym usunięciu węzła przeliczane były ponownie stopnie wszystkich pozostałych. W sytuacji, gdy występowały centra o tym samym stopniu – jedno z nich było losowane z rozkładem równomiernym.

Pozostaje kwestia otwartą, czy eliminacja centrów jest najbardziej niszczycielska w stosunku do sieci. W każdym ruchu usunięte zostaje maksymalnie wiele krawędzi. Jednak możliwe, że usuwanie wierzchołków, np. w celu większej fragmentacji sieci, powinno następować według innych reguł.

4. GENEROWANIE SIECI

Generowaniu i klasyfikacji sieci bezskalowych poświęcono już szereg publikacji [2], [8]. Różnorodność opisywanych sieci bezskalowych wynika z faktu istnienia wielu dziedzin nauki, w których można zaobserwować takie topologie. Zanim jednak opisane zostaną poszczególne modele generowania sieci bezskalowych, warto przytoczyć obserwacje dwóch uczonych, którzy jako pierwsi racjonalnie wytłumaczyli, dlaczego Internet ma topologię sieci bezskalowej.

Watts i Strogatz [9] sformułowali dwa postulaty przesądzające o tym, że wynikowa sieć będzie miała charakter bezskalowy. Pierwszy postulat głosi, że sieć rozrasta się. Oznacza to, że węzłów w sieci przybywa, czyli trudniej zamodelować stan obecny niż przewidywać rozrost sieci od załóżka. Postulat drugi jest mniej oczywisty – węzły podłączają się do istniejącej struktury preferując jako miejsce podłączenia już istniejące centra. Oczywiście nie jest to deterministyczne, wybór węzła jest losowy. Zmieniają się tylko parametry samego losowania. Najczęściej preferencja ma charakter liniowy, co oznacza, że prawdopodobieństwo podłączenia do węzła posiadającego osiem krawędzi jest dwukrotnie wyższe niż do węzła posiadającego cztery. Powyższe dwa postulaty pozwalają wytłumaczyć dzisiejszą topologię sieci Internet.

4.1. Parametry sieci

Badania przez Watts'a i Strogatza szeregu istniejących sieci pozwoliły im wysnuć wniosek o pośredniej naturze sieci bezskalowych. Regularne sieci, jak np. siatka (ang. *lattice*), cechuje wysoki współczynnik grupowania CC (ang. *clustering coefficient*). Sieci losowe charakteryzuje niewielka średnia długość najkrótszych ścieżek pomiędzy wierzchołkami – charakterystyczna długość ścieżki. Prawidłowości w strukturze sieci bezskalowych powodują, że przy zachowaniu losowości łączą obie powyższe cechy. Charakterystyczna długość ścieżki jest mniejsza niż w sieciach regularnych, a współczynnik grupowania wyższy.

Podstawową cechą sieci bezskalowych jest potęgowy rozkład liczby węzłów w zależności od stopnia węzła. Rozkład ten przesądza o zaklasyfikowaniu sieci do kategorii bezskalowych. Rozkład ten powoduje, że występują w sieci centra, czyli węzły posiadające dużą, w stosunku do pozostałych węzłów, liczbę krawędzi. Histogram stopni węzłów w podwójnej skali logarytmicznej tworzy w przybliżeniu prostą (rys. 1).

Nachylenie tej prostej, oznaczane przez² α (< 0), będące wykładnikiem potęgi, jest parametrem analizowanej sieci bezskalowej.

Drugą cechą służącą charakteryzowaniu sieci jest współczynnik CC . Współczynnik ten jest wartością średnią współczynników wyznaczonych dla wszystkich węzłów sieci. Konkretny wierzchołek o numerze i posiada sąsiadów w liczbie równej własnemu stopniowi. Sami sąsiedzi, bez i -tego wierzchołka, wraz z połączeniami między nimi tworzą graf G_i . Współczynnik grupowania określa stopień, w jakim połączony jest graf G_i w stosunku do grafu pełnego:

$$CC_i = \frac{\text{liczba_krawedzi_grafu_}G_i}{k_i(k_i - 1)/2} \quad (4.1)$$

oraz:

$$CC = \frac{1}{N} \sum_i CC_i \quad (4.2)$$

gdzie N jest liczbą węzłów całej sieci.

Jak wspomniano, kolejnym kryterium określającym parametry sieci jest charakterystyczna długość ścieżki – L (ang. *characteristic path length*). Długość ta jest średnią długości najkrótszych ścieżek d_{ij} (ang. *distance*) pomiędzy wszystkimi parami wierzchołków występującymi w grafie:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \quad (4.3)$$

Wymienione powyżej parametry sieci zostaną wykorzystane do zobrazowania różnic w reakcji sieci losowych i bezskalowych na uszkodzenia.

4.2. Model Watts'a – Strogatza (WS)

Model ten [9] zaproponowany został na podstawie obserwacji natury szeregu sieci bezskalowych. Ponieważ naukowcy doszli do wniosku, że sieci bezskalowe są pośrednio sieciami regularnymi, generację sieci N -węzłowej rozpoczyna się od wygenerowania

² często także γ , $\gamma = -\alpha$;

regularnej siatki. Następnie z prawdopodobieństwem p , dla każdej krawędzi, zmieniają się wierzchołki, które krawędź łączyła.

4.3. Model Barabasi'ego – Alberta (BA)

Jest to jeden z bardziej popularnych, a zarazem prostych w implementacji modeli generowania sieci bezskalowych, zaproponowany w [10]. Jest to model dynamiczny. Oznacza to, że sieć ma początkowo mniejszą liczbę wierzchołków niż sieć docelowa. Jak wspomniano, model ten korzysta z dwóch postulatów: rozrostu sieci i preferencyjnego przyłączania. Algorytm generowania sieci o N węzłach przedstawia się następująco:

- a) zaczynając od $m_0 \ll N$ węzłów w każdej chwili czasu dodaj węzeł o m ($\leq m_0$) krawędziach
- b) prawdopodobieństwo przyłączenia krawędzi nowego węzła do już istniejącego i -tego węzła o stopniu k_i jest równe ilorazowi stopnia k_i i sumy wszystkich stopni (4.4).

$$P(k_i) = \frac{k_i}{\sum_j k_j} \quad (4.4)$$

Model ten nie pozwala na generowanie sieci z większym CC , ani też dowolny wybór średniego stopnia K węzłów sieci. Dla modelu BA zachodzi zależność (4.5).

$$K \cong 2 \cdot m \quad (4.5)$$

Oznacza to, że średni stopień wierzchołków grafu jest równy (z dokładnością do $(2K_0(m-1))/N$) podwojonej liczbie krawędzi dodawanych z każdym nowym węzłem.

4.4. Model Generalized Linear Preference (GLP)

Model uogólnionej preferencji liniowej GLP, zaproponowany w [8] stanowi połączenie idei generatorów WS i BA. Jest to także model dynamiczny, w którym rozpoczyna się od $m_0 \ll N$ węzłów i stopniowo dodawane są kolejne węzły. W każdym kroku zachodzi ono jednak z pewnym prawdopodobieństwem p . Oprócz dodawania węzłów, w każdym kroku, z prawdopodobieństwem $1-p$, zachodzi także dodawanie m połączeń pomiędzy już istniejącymi wierzchołkami. Stosowana jest także preferencja liniowa. Proces ten ma modelować zachowanie dostawców Internetu – operatorów domen – polegające na rozbudowie posiadanych łączy o nowe w celu rozszerzenia pasma. Model GLP jest szczególnie ukierunkowany na uzyskiwanie dowolnej wartości parametru α sieci bezskalowych.

4.5. Model kształtowania średniego stopnia sieci

Proponowany tutaj model pozwala dowolnie kształtować średni stopień sieci K . Model ten jest modyfikacją modelu BA, czyli zalicza się również do modeli dynamicznych. Różnicą w stosunku do omówionego modelu BA jest fakt losowania liczby krawędzi, jaką ma nowy wierzchołek. Losowanie odbywa się ze zmodyfikowanym (odrzucając wartości poniżej 1) rozkładem normalnym wokół docelowego średniego stopnia K wierzchołków grafu wynikowego.

5. SPRAWNOŚĆ SIECI

Jako miara sprawności sieci zostały wybrane trzy kryteria: współczynnik grupowania CC , długość charakterystycznej ścieżki L i współczynnik ubytku łączy U . Współczynnik ubytku łączy jest proponowanym kryterium oceniania sprawności sieci. Pozwala on na łatwiejsze porównanie sieci między sobą. Jest on średnią współczynników dla wszystkich węzłów sieci (5.2). Dla każdej chwili t po usunięciu węzłów z sieci wartość U dla danego węzła i jest unormowaną liczbą krawędzi, jakie aktualnie on posiada (5.1).

$$U_i(t) = \frac{k_i(t)}{k_i(t=0)} \quad (5.1)$$

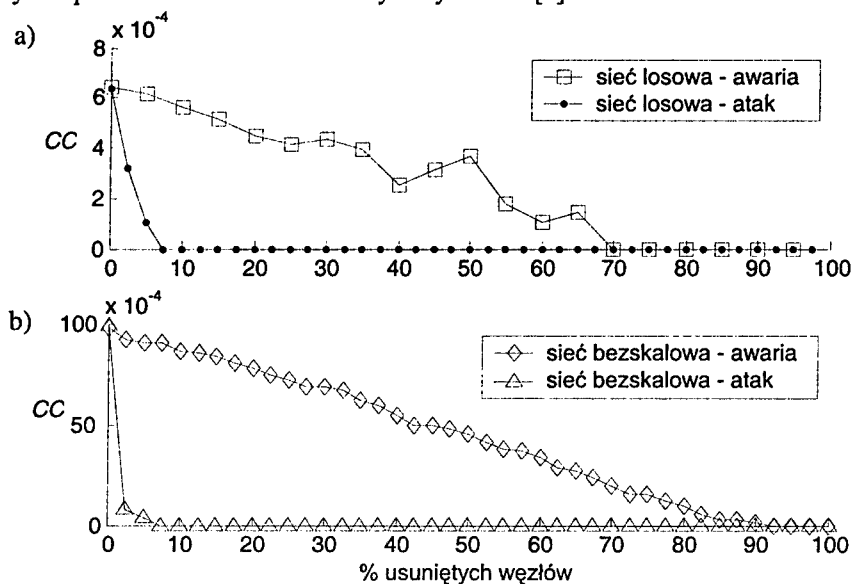
oraz:

$$U(t) = \frac{1}{N} \sum_i U_i(t) \quad (5.2)$$

gdzie N jest liczbą wszystkich węzłów sieci.

6. WYNIKI BADAŃ

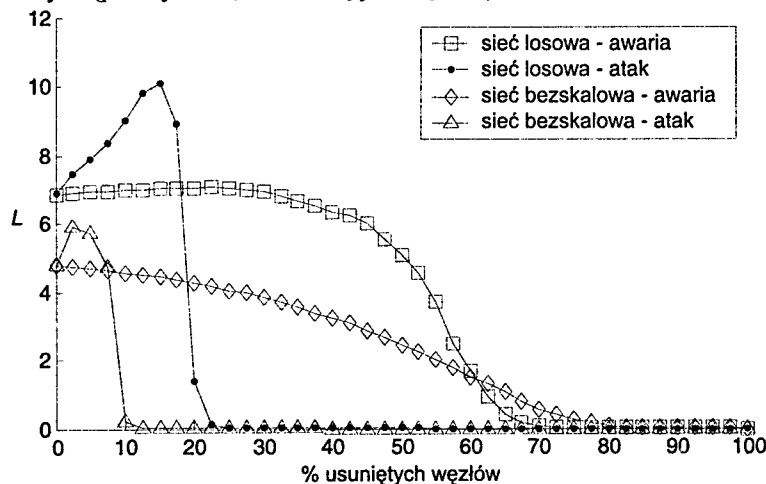
Prezentowane wyniki są rezultatami symulacji pięciu sieci bezskalnych i pięciu losowych o $N=2000$ i $K=3$, przy czym każdy scenariusz awarii lub ataku powtarzany był dziesięciokrotnie. Wartość współczynnika γ dla sieci bezskalnych miała wartość rzędu 2.2, czyli odpowiadała wartościom rzeczywistych sieci [4].



Rys.2. Współczynnik grupowania; a) – sieć losowa, b) – sieć bezskalowa

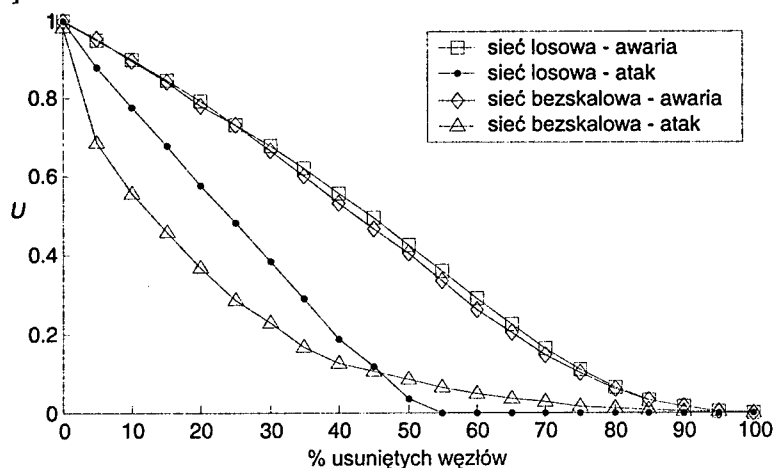
Wyniki symulacji przedstawione na rys. 2 pokazują, że badane sieci bezskalowe rzeczywiście posiadały wyższą wartość współczynnika grupowania. Wykreślone zostały osobno charakterystyki dla sieci bezskalnych i losowych. Można dzięki temu zaobserwo-

wać, że w atakowanych sieciach bezskalowych szybciej procentowo spada wartość CC niż w sieciach losowych (punkty dla 2,5% usuniętych węzłów).



Rys. 3. Charakterystyczna długość ścieżki

Charakterystyczna długość ścieżki potwierdza, że wielkość tego parametru dla sieci bezskalowych jest mniejsza niż dla losowych (rys. 3). Interesujące jest szybsze narastanie L w przypadku atakowania sieci bezskalowych. Oznacza to, że eliminacja węzłów nie powoduje jeszcze rozpadu sieci na mniejsze, lecz wydłużenie tras – szybsze w sieciach bezskalowych. Po osiągnięciu maksimum charakterystyki gwałtownie spadają, co jest oznaką rozpadania centralnej grupy węzłów na mniejsze, z czym związane jest zjawisko perkolacji [11].



Rys. 4. Ubytek łączy

Ostatnie kryterium pozwalające zaobserwować reakcje sieci na uszkodzenia to współczynnik ubytku łączy (rys. 4). Na wykresie zobrazowanych jest kilka ciekawych cech sieci bezskalowych. Atakowane sieci losowe szybciej osiągają stadium zerowej sprawności.

Wynika to z występowania w tych sieciach węzłów o stosunkowo wysokim stopniu, około wartości średniej. W sieciach bezskalowych po usunięciu centrów pozostałe węzły mają niższy stopień niż przeciętne węzły sieci losowych. Dlatego sprawność sieci losowych szybciej osiąga zero. Jednocześnie sprawność sieci losowych jest wyższa w środku przedziału, zarówno dla awarii jak i losowego ataku.

7. PODSUMOWANIE

Przeprowadzone symulacje ataków i awarii ujawniają podatność sieci bezskalowych na ataki w dużo większym stopniu niż sieci losowych. Metoda eliminacji węzłów ma większe znaczenie w sieciach bezskalowych niż w sieciach losowych. Zaproponowany współczynnik ubytku łączy pozwala łatwo porównywać stopień sprawności różnorodnych sieci. Prezentowane wyniki badań prezentują, jak szybko może zostać zachwiana funkcjonalność realnych sieci bezskalowych, której przykładem jest Internet. Otwartą kwestią pozostaje pytanie, czy eliminacja centrów jest optymalnym sposobem ataku.

BIBLIOGRAFIA

- [1] Erdős P., Rényi A.: *On the Evolution of Random Graphs*, Publ. Math. Inst. Hung. Acad. Sci. 1960
- [2] Goh K.-I., Oh E., Jeong H., Kahng B., Kim D.: *Classification of scale-free networks*, Proc. Natl. Acad. Sci. U.S.A. 99, 12583-12588 (2002).
- [3] A.-L. Barabási, E. Bonabeau: *Sieci bezskalowe*, Świat Nauki, czerwiec 2003
- [4] M. Faloutsos, P. Faloutsos, C. Faloutsos: *On Power-Law Relationships of the Internet Topology*, Proc. of ACM SIGCOMM'99, Aug. 1999
- [5] P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda: *Efficiency of Scale-Free Networks: Error and Attack Tolerance*, cond-mat/0205601, May 2002
- [6] University of Oregon Route Views Project <http://www.antc.uoregon.edu/route-views/>
- [7] Routing Information Service <http://www.ripe.net/ripenc/pb-services/np/ris/index.html>
- [8] T. Bu, D. Towsley: *On Distinguishing between Internet Power Law Topology Generators*, IEEE Infocom 2002
- [9] D. J. Watts and S. H. Strogatz: *Collective dynamics of "small-world" networks*. Nature, 393:440-442, 1998
- [10] A.-L. Barabási, R. Albert, *Emergence of scaling in random networks*, Science, Oct. 1999
- [11] D.-S. Lee, K.-I. Goh, B. Kahng, D. Kim: *On the evolution of scale-free graphs*, cond-mat/0312336, Dec 2003

RESISTANCE TO FAILURES OF SCALE-FREE NETWORKS

Summary

This paper discusses differences between random and scale-free networks in the aspect of node elimination. Two models of elimination are presented: random failure and intentional attack. The differences in results for random and scale-free networks are commented. Simulation results are presented using various network functionality criteria. A new criterion, allowing to compare different types of networks, is introduced.

Grzegorz Górski

Katedry Systemów Informacyjnych, Politechnika Gdańska

DYNAMICZNE SIECI VLAN WYKORZYSTUJĄCE INFRASTRUKTURĘ KLUCZA PUBLICZNEGO PKI

Streszczenie

W artykule dokonano analizy bezpieczeństwa przesyłania danych w „dynamicznej” sieci VLAN za autoryzacją dostępu. Zaproponowano bezpieczny protokół autoryzacji dostępu użytkownika do zasobów sieciowych. Zakłada on stosowanie metod ochrony przy przesyłaniu danych pomiędzy wszystkimi komponentami algorytmu. Kluczowym elementem jest użycie infrastruktury klucza publicznego PKI, pozwalającej na znaczący wzrost bezpieczeństwa przesyłanych informacji w „grupie wirtualnej”. Mechanizmy wywodzące się ze standardu X.509, wykorzystujące kryptografię z kluczami niesymetrycznymi są szczególnie użyteczne w wirtualnych sieciach dynamicznych. Przejmuje się przy tym, że nowy użytkownik takiej domeny rozgłoszeniowej nie posiada żadnych informacji na temat używanych w danej chwili, w grupie, kluczy szyfrujących. Po zweryfikowaniu tożsamości użytkownika to właśnie infrastruktura klucza publicznego pozwala na bezpieczne przekazanie kluczy.

1. WSTĘP

Wirtualną lokalną sieć komputerową VLAN można zdefiniować [1] jako grupę użytkowników (stacji), należących do różnych segmentów sieci, którzy mogą komunikować się tak, jakby należeli do jednej domeny rozgłoszeniowej (podsieci lokalnej). Istnieją dwa rodzaje wirtualnych grup roboczych. Pierwsza określana jest mianem statycznej sieci VLAN [2]. Sieci tego rodzaju zbudowane są według zasady, w której administrator statycznie (w trybie off-line) przydziela wybraną stację roboczą do danej grupy wirtualnej. Korzystając z informacji o adresach stacji, które zawarte są w napływających od stacji nagłówkach pakietów, administrator tworzy listę grup roboczych, do których powinna należeć dana stacja robocza. Drugi rodzaj sieci VLAN to tzw. dynamiczne grupy wirtualne. Ten rodzaj protokołów sieciowych pojawił się dopiero kilka lat temu, kiedy przełączniki sieciowe zaczęły przetwarzać informacje wstawiane do pakietów przez trzecią i wyższe warstwy protokołów (według modelu ISO/OSI).

W dynamicznej sieci VLAN zakłada się, że liczba komputerów w grupie może się zmieniać w czasie. Z tego powodu algorytmy dynamicznych sieci VLAN muszą dodatkowo obsługiwać procesy podłączania i odłączania stacji od sieci. Protokoły te nie wymagają żadnej interwencji administratora, dlatego są nazywane AUTO-VLANs. Przyna-

leżność do wybranej, dynamicznej grupy może zależeć od typu aplikacji, usług uruchamianych przez użytkownika (np. usługa poczty internetowej dostępna pod określonym standardowo portem protokołu TCP) lub identyfikatora użytkownika.

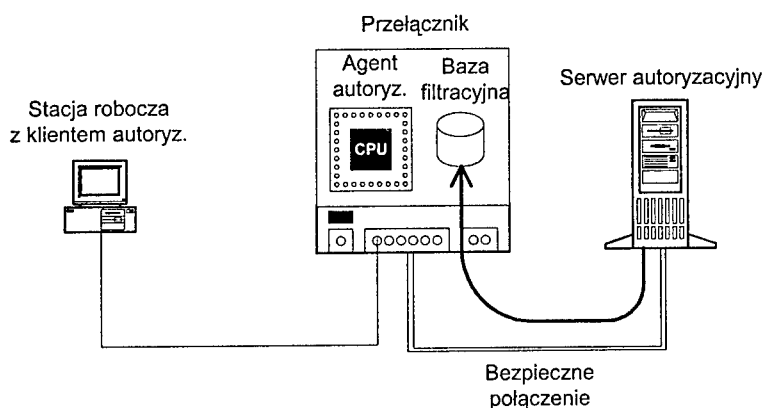
2. DYNAMICZNE SIECI VLAN Z AUTORYZACJĄ DOSTĘPU

Sieci VLAN z autoryzacją dostępu [3] są dobrym przykładem dynamicznych wirtualnych grup roboczych. Proces autoryzacji może opierać się na trzech możliwych elementach: porcie fizycznym, regule logicznej oraz identyfikatorze użytkownika. Weryfikacja użytkownika oparta o fizyczny port zakłada, że administrator definiuje wybrany port przełącznika jako zaufany dla wybranej wirtualnej grupy roboczej. Na tej podstawie cały ruch pakietów pochodzący z tego portu jest uznawany za autoryzowany. Oczywiście to proste rozwiązanie nie zapewnia wymaganego wysokiego poziomu bezpieczeństwa. Niedogodność tą usunięto w algorytmie z regułą logiczną, w którym port przełącznika jest powiązany z listą adresów sieciowych użytkowników. Tylko koniunkcja portu przełącznika z odpowiednim adresem sieciowym użytkownika zapewnia poprawną autoryzację stacji roboczej. Identyfikator użytkownika jest ostatnim czynnikiem na podstawie może być wykonany proces autoryzacji. Metoda ta pozwala na najbezpieczniejszy, w stosunku do dwóch opisanych powyżej, dostęp do zasobów sieciowych. Powoduje to jednak, że odpowiedni protokół obsługujący taki typ dynamicznego VLAN'u jest najbardziej złożony.

Wirtualna sieć LAN z weryfikacją dostępu [4] składa się z trzech niezależnych komponentów. Serwer autoryzacyjny przechowuje profile użytkowników (zestawy te zawierają: identyfikator użytkownika, hasło, metodę weryfikacji, listę dostępnych zasobów przedziały czasu, kiedy użytkownik może się logować do grupy roboczej i do przypisanych mu zasobów). Serwer musi także posiadać listę adresów wszystkich bezpośrednio przyłączonych do niego przełączników oraz właściwe zestawy kluczy kryptograficznych do zestawienia bezpiecznych połączeń do wymiany poufnych informacji pomiędzy komponentami całego algorytmu VLAN. Kolejnym elementem opisywanego algorytmu dynamicznych grup roboczych jest agent autoryzacyjny. Jest on instalowany jako specjalizowane oprogramowanie dla systemu operacyjnego przełącznika sieciowego. Implementacji tego składnika silnie zależy od wybranego modelu urządzenia i jego oprogramowania (ang. firmware). Jednakże z tego powodu oprogramowanie agenta może być używane w heterogenicznych środowiskach sieciowych. Agent autoryzacyjny pracuje w dwóch trybach określanych jako tryb czuwania w przypadku gdy nie ma żadnej aktywnej grupy roboczej oraz w trybie aktywnym w przeciwnym wypadku. Przy każdorazowym uruchamianiu oprogramowania agenta tworzony jest bezpieczny kanał transmisji danych do głównego serwera autoryzacyjnego. Ze względów bezpieczeństwa każdy agent posiada listę dodatkowych serwerów autoryzacyjnych w przypadku awarii głównego serwera. Taka konfiguracja wymaga niezależnego zdefiniowania pary adresu IP oraz klucz autoryzacyjnego dla algorytmu szyfrującego dla każdego dostępnego serwera autoryzacyjnego. Agent jest centralnym komponentem całego algorytmu, który bierze udział w procesie weryfikacji i ze względów bezpieczeństwa oddziela serwer autoryzacyjny i oprogramowanie klientów. Wszystkie żądania dostępu do zasobów muszą być przetworzone przez agenta autoryzacyjnego, z tego powodu dodatkowy program monitorujący jego poprawną pracę musi być dodatkowo zaimplementowany w przełączniku. Zlicza on liczbę nieudanych połączeń użytkowników i wysyła komunikaty do administratora systemu. Ostatnim komponentem prezentowanego algorytmu VLAN jest klienta autoryzacyjny. Klient ten to specjalizowane oprogramowanie zainstalowane na komputerze PC w celu

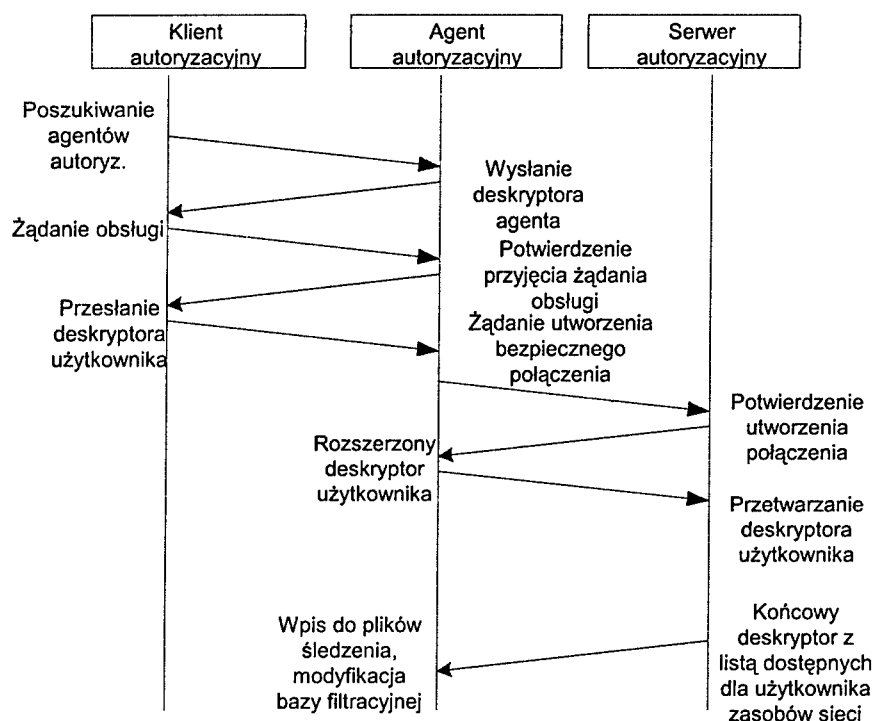
komunikacji z agentami autoryzacyjnymi. Zadaniem tego składnika jest wykryć adres sieciowy dostępnych agentów, dlatego właśnie proces logowania z reguły rozpoczyna się od pakietu rozgłoszeniowego wysyłanego ze stacji roboczej. Takie rozwiązania nie wymaga odrębnej konfiguracji dla każdego komputera w sieci.

Pracę opisywanego algorytmu dynamicznej sieci VLAN z autoryzacją dostępu można podzielić na trzy zasadnicze etapy. Dostarczenie żądania połączenia do sieci inicjuje pierwszy etap. Klient autoryzacyjny wykrywa zestaw adresów IP dostępnych w sieci przełączników z zainstalowanym oprogramowaniem agenta oraz wysyła deskryptor autoryzacyjny do agenta z wykorzystaniem właściwego dla tej usługi portu protokołu TCP. Etap ten został zaprezentowany na rysunku 1.



Rys.1. Pierwszy etap procesu autoryzacji – żądanie połączenia do sieci.

W kolejnym etapie agent autoryzacyjny przetwarza przychodzące żądanie obsługi. Najpierw, dodaje do deskryptora użytkownika nowe informacje takie jak: czas zgłoszenia, port przełącznika i adres sieciowy użytkownika a następnie przesyła te dane przez poprzednio utworzony bezpieczny kanał do serwera autoryzacyjnego. Na podstawie informacji zawartych w deskryptorze takich jak: identyfikator użytkownika, hasło, adres sieciowy, czas systemowy serwer decyduje o przyznaniu dostępu do sieci użytkownikowi. Jest to trzeci i najbardziej istotny etap całego procesu weryfikacji. W przypadku pozytywnej autoryzacji serwer dodaje do deskryptora listę wraz przedziałami czasu dostępnych dla użytkownika zasobów, a następnie odsyła tę informację z powrotem do właściwego agenta autoryzacyjnego. Na podstawie pełnego deskryptora użytkownika agent autoryzacyjny zapisuje informacje dotyczącą wyniku procesu weryfikacji do pliku dziennika. Czynność ta rozpoczyna ostatni etap procesu autoryzacji określany jako modyfikacja bazy filtrującej ruch pakietów przełącznika. Na podstawie listy dostępnych zasobów otrzymanych z serwera agent „kojarzy” zasoby z ich siecią lokalizacją. Taka tablica użytkownika jest częścią ogólnej bazy danych zasobów sieci, która jest obsługiwana przez wybranego agenta. Tablica klienta jest wykorzystywana przez algorytmy systemu operacyjnego przełącznika w celu kontroli ruchu pakietów przychodzących i wychodzących do podłączonego użytkownika. Pełny proces autoryzacji dostępu użytkownika do zasobu sieci pokazano w postaci diagramu na rysunku 2.



Rys.2. Proces autoryzacji dostępu użytkownika do zasobu.

Protokoły dynamicznych sieci VLAN muszą także obsługiwać proces odłączania stacji od sieci. Odłączenie stacji od zasobu sieciowego jest wykonywane zwykle, gdy agent otrzyma od klienta żądanie odłączenia. Można jednakże określić kilka scenariuszy błędnych rozłączeń sesji. Jeden z nich odpowiada przypadkowi, gdy przełącznik wykrywa brak sygnału na porcie fizycznym użytkownika i odłącza stację od autoryzowanej grupy VLAN. Innymi powodami przerwania połączenia może być zbyt długi czas bezczynności stacji, zbyt długi czas dostępu do zasobu (dłuższy niż określono w deskryptorze z serwera autoryzacyjnego). Informacja ta jest przesyłana od serwera do agenta w trakcie trzeciego etapu procesu autoryzacji. W celu poprawy poziomu bezpieczeństwa wszystkie pakiety przychodzące od klientów (adresów sieciowych) innych niż te zarejestrowane jako zweryfikowane pozytywnie powinny być odrzucane na portach przełącznika.

3. BEZPIECZEŃSTWO PRZESYŁANIA DANYCH W DYNAMICZNYCH SIECIACH VLAN Z AUTORYZACJĄ DOSTĘPU

Bezpieczeństwo pracy dynamicznych sieci VLAN z autoryzacją dostępu silnie zależy od sposobu, w jaki poszczególne komponenty algorytmu wymieniają między sobą informacje poufne. Prezentowany powyżej algorytm zakłada, że do komunikacji pomiędzy agentem i serwerem autoryzacyjnym używany jest bezpieczny kanał i wszystkie informacje są szyfrowane. Zabezpieczenia te dotyczą jednak warstwy sieciowej (np. IPsec) lub warstw wyższych modelu ISO/OSI. Dokładniejsza analiza [5] pokazuje, że najbardziej narażonym

na atak jest oprogramowanie agenta autoryzacyjnego. Atak taki wykorzystując informacje zawarte w warstwie łącza danych może spowodować przejęcie kontroli nad przełącznikiem i kierowaniem strumieni pakietów. W takim wypadku włamywacz może podłączyć się do zasobu sieci, do którego nie ma uprawnień lub uzyskać niepowołany dostęp do informacji wymienianych wewnątrz grupy. Pozostali członkowie grupy wymieniający między sobą dane nie posiadają o tym jednak żadnej informacji.

Jednym z najbardziej znanych ataków jest przepełnienie tablic adresowych przełącznika (ang. *MAC flooding*). Każdy przełącznik przetwarza w swojej pamięci tablicę kojarzącą adresy MAC obsługiwane na poszczególnych portach urządzenia z numerami sieci VLAN, a także innymi parametrami. Pojawienie się ramki o nieznanym adresie warstwy łącza danych powoduje stworzenie nowego wpisu w tablicy. Ze względu na skończoną pamięć przełącznika ograniczone są także maksymalne wymiary tablicy. Przeprowadzenie ataku jest proste wiąże się ono z wysyłaniem do komutatora wielkiej liczby ramek o różnych adresach MAC. Z powodu zapełnienia tablicy urządzenie przestaje uaktualniać topologię otaczającej sieci i wysyła każdą ramkę typu unicast zamiast do jednego portu do wszystkich należących do tej samej sieci VLAN (domeny rozgłoszeniowej). Powoduje to rozchodzenie się fali nadmiarowych pakietów na inne przełączniki w sieci, zmniejszenie wydajności pracy, ale przede wszystkim umożliwia włamywaczowi obserwację ruchu w sieci. Tego typu zagrożenie można ograniczyć wprowadzając limit adresów MAC dla wybranego portu urządzenia, powoduje to jednak ograniczenie możliwości komunikowania się zwłaszcza w dużych sieciach korporacyjnych.

Zagrożeniem bezpieczeństwa stosowania sieci VLAN mogą być także protokoły trunkingowe służące do przesyłania pomiędzy komutatorami informacji o przynależności do wirtualnych grup roboczych. Nie jest możliwe ograniczenie stosowania protokołów tego typu, gdyż ograniczyłoby to sieć VLAN praktycznie do pojedynczego przełącznika. Protokoły trunkingowe np. zgodne ze standardem 802.1q [1] czy też Cisco ISL powinny być uruchamiane na portach, do których są podłączone wyłącznie znane przełączniki lub inne systemy wymagające zestawienia tego typu magistrali. W przypadku ustawienia portu dostępowego (statycznie lub dynamicznie) w sposób umożliwiający zestawienie traktu 802.1q, umożliwia się włamywaczowi wysyłanie ramek do wszystkich sieci VLAN w ramach magistrali (tzw. VLAN leaking). Dużo większe zagrożenie niesie atak typu VLAN-hopping wykorzystujący implementacje tzw. Native VLAN w ramach protokołu 802.1q umożliwiające wykorzystanie starszych typów przełączników. Rozszerzenie to umożliwia transmisję z wykorzystaniem traktu ramek bez znacznika określającego numer sieci wirtualnej. Nawet ramki opatrzone znacznikiem są przed wysłaniem jego pozbawiane, co daje włamywaczowi możliwość zaatakowania, poprzez przełącznik obsługujący Native VLAN, dowolnej sieci VLAN. Ramki ze znacznikiem 802.1q kierowane do Native VLAN są obsługiwane także na portach, na których protokół 802.1q jest wyłączony (typowe porty dostępowe). Atak tego typu działa w jednym kierunku jest jednak wykorzystywany do realizacji skutecznych ataków typu DoS (odmowa dostępu) wewnątrz sieci. Prawdopodobieństwo tego ataku można ograniczyć wyłączając protokoły na wszystkich portach, na których nie są stosowane tzn. w podsieciach gdzie nie ma członków danych grup roboczych. Ponadto zaleca się wyłączenie z obsługi tych podsieci wirtualnych, które nie są wykorzystywane w danym obszarze obsługiwanym przez wybraną magistralę. Pomimo możliwości wykorzystania urządzeń aktywnych starszych generacji, jakie oferuje implementacja protokołu Native VLAN, ze względów bezpieczeństwa danych przesyłanych w sieci rozwiązanie to należy traktować jako przejściowe. Docelowo zaś we wszystkich domenach rozgłoszeniowych powinien być aktywowany proces znacznikowania ramek.

Kolejnym protokołem działającym w sieciach komutowanych a wykorzystywanym do ataków może być ARP (ang. *Address Resolution Protocol*), który odpowiada za odwzorowanie adresów warstwy sieciowej np. IP na adresy warstwy drugiej MAC. Włamywacz może samorzutnie wysyłać w formie rozgłoszeniowej odpowiedzi bez zaproszenia, czyli tzw. Gratuitous ARP, które są generowane przez systemy np. po restarcie lub zmianie adresu sieciowego. Powoduje to, że inne urządzenia sieciowe posiadające w tablicach ARP Cache pary np. MAC-IP zaktualizują swoje wpisy. Protokół ten nie wymaga żadnego uwierzytelnienia, co powoduje, że włamywacz może podszyć się pod system i nawet przechwycić całość komunikacji w danej sieci VLAN lub do wybranej stacji roboczej. Są to ataki tzw. MIM (ang. *Man in the Middle*), w których włamywacz działa jako pośrednik kontrolując wymianę informacji od innych użytkowników. Protokół ARP może być także wykorzystany do ataku typu ARP Spoofing, w którym włamywacz stara się przechwycić ruch generowany przez stacje robocze do routera. Dopiero zastosowanie rozwiązania ARP Inspection polegającego na przypisaniu pojedynczego adresu MAC oraz funkcjonalności typu Private VLAN zmniejszyło zagrożenie tego typu atakiem. Rozwiązanie Private VLAN polega na tworzeniu takich grup wirtualnych, z którym możliwe jest kierowanie ruchu do portów ogólnodostępnych, do których podłączone są serwery i routery. Ruch pomiędzy portami dostępowymi jest sprzętowo blokowany, przez co podszycie się przy użyciu ARP Spoofing pod innego użytkownika w sieci VLAN staje się niemożliwe dzięki zasadzie Sticky ARP uniemożliwiającej przełącznikowi dokonanie nowego lub korektę dla istniejącej już pary MAC-IP. Zmiana wymaga ręcznej interwencji administratora.

4. ZASTOSOWANIE INFRASTRUKURY PKI W SIECIACH AUTO-VLAN

Przedstawione w poprzednim rozdziale możliwości zaatakowania sieci przełączanych wykorzystujących algorytmy dynamicznych sieci VLAN pokazują, że zastosowanie zabezpieczonych kanałów transmisji pomiędzy agentem i serwerem autoryzacyjnym nie rozwiązuje problemu zapewnienia bezpieczeństwa. Włamanie, a dokładniej biorąc przejęcie kontroli nad strumieniem pakietów może nastąpić pomiędzy klientem i serwerem autoryzacyjnym. Przedstawiony w rozdziale 2 algorytm sieci AUTO-VLAN zakłada, że agenci posiadają listę aktywnych serwerów oraz, że liczba agentów jest niewielka i ograniczona geograficznie. Takie założenia nie są możliwe do spełnienia w sieci korporacyjnej, w której trudno byłoby wyobrazić sobie powtarzaną okresowo, ręczną procedurę wymiany kluczy szyfrujących dla agentów i serwerów autoryzacyjnych.

Rozwiązaniem tego problemu jest zastosowanie infrastruktury klucza publicznego PKI jako mechanizmu służącego do wzajemnego uwierzytelniania się poszczególnych komponentów przedstawionego powyżej protokołu. Rozwiązanie to nie zakłada posiadania żadnej wiedzy przez użytkownika [6] przed rozpoczęciem procedury autoryzacji. Jedynym wymaganiem jest wyposażenie wszystkich użytkowników oraz agentów i serwerów autoryzacyjnych w aktywne certyfikaty zgodne ze standardem X.509.

Przy rozpoczęciu procedury podłączania komponent klienta sprawdza na serwerze lub serwerach autoryzacyjnych listę aktywnych agentów oraz ich certyfikaty w celu pobrania klucza publicznego jednego z nich. Następnie wysyła do wybranego agenta żądanie podłączenia do sieci szyfrując przesyłane dane jego publicznym, sprawdzając uprzednio czy certyfikat agenta jest ważny tzn. nie został opublikowany na listach CRL lub przy użyciu protokołu OCSP (ang. *On-line Certificate Status Protocol*). Tylko wybrany agent będzie mógł przy użyciu swojego klucza prywatnego rozszyfrować skierowaną do niego wiado-

mość. Następnie wysłany jest do użytkownika aktywny, w danej chwili, klucz sesyjny umożliwiający przesyłanie zaszyfrowanych informacji do innych członków grupy wirtualnej. Podobnie, proces uwierzytelnienia musi zostać wykonany dla każdego agenta z właściwym serwerem autoryzacyjnym, który po zakończonym podłączeniu agenta wpisuje go na listę aktywnych agentów w sieci. Wykorzystanie mechanizmów PKI zostało ze względów wydajnościowych ograniczone do bezpiecznego przekazania symetrycznych kluczy sesyjnych. Dalszy wzrost poziomu bezpieczeństwa przesyłanych w grupie danych można osiągnąć wymieniając okresowo klucze sesji. Proces wymiany klucza jest inicjalizowany przez serwer autoryzacyjny, który rozsyła do wybranych agentów nowy klucz szyfrujący grupy. Samo przekazanie nowego klucza może być wykonane w formie przesłania do użytkownika zakodowanego nowego klucza z wykorzystaniem starego klucza sesji jako klucza szyfrującego. Do przekazania nowego klucza można też użyć ponownie klucza publicznego użytkownika. To drugie rozwiązanie jest powieleniem bardziej złożonego obliczeniowo etapu autoryzacji, ale jest bezpieczniejsze, a ponieważ klucze są wymieniane okresowo np. co 5 minut nie powoduje to znacznego spadku efektywności działania protokołu.

Rozwiązanie z zastosowaniem certyfikatów standardu X.509 pozwala na bezpieczne podłączanie nowych użytkowników. Sam proces logowania członka grupy może być oparty na identyfikatorze i hasle dla danej grupy roboczej, które użytkownik musi znać, lub też tylko na identyfikacji użytkownika na podstawie certyfikatu. W takim przypadku administrator systemu musi przygotować na serwerze autoryzacyjnym profil każdego użytkownika z listą wirtualnych grup roboczych, do których należy. Obie metody administrowania pozwalają jednak na tworzenie bezpiecznych dynamicznych sieci VLAN z automatyczną dostępu, w których liczba członków grupy może się zmieniać w czasie.

5. ZAKOŃCZENIE

Powszechnym wymaganiem stawianym przed administratorami sieci komputerowych w globalnych korporacjach jest nie tylko zabezpieczenie sieci wewnętrznej przed atakami z zewnątrz, ale także możliwość podziału użytkowników na grupy robocze, których członkowie wymieniają poufne, niedostępne dla innych pracowników informacje. Charakter ruchu generowanego przez członków takiej grupy może być bardzo różny w zależności od tego czy jest to bezpieczna wideokonferencja czy rozmowa poprzez tekstowy komunikator sieciowy. Podstawowe wymaganie dotyczy zawsze gwarantowanego poziomu zabezpieczenia oraz możliwości śledzenia dynamicznego procesu podłączania i odłączania użytkowników. Przedstawiona w artykule analiza zagrożeń, utrudnienia związane z dystrybucją dużej liczby kluczy szyfrujących w sieci korporacyjnej jednoznacznie wskazują, że dynamiczne sieci VLAN wymagają zastosowania technologii PKI. Rozwiązanie to nie zakłada posiadania przez użytkownika żadnej wiedzy o kluczach szyfrujących używanych w wybranej grupie wirtualnej. Jedynym wymaganiem jest wyposażenie użytkowników w certyfikaty klucza publicznego zgodne ze standardem X.509. Badania wydajności poszczególnych algorytmów kryptograficznych pokazują jednak, że infrastruktura PKI jest wykorzystywana tylko do bezpiecznego dostarczenia aktualnego w danej chwili w grupie symetrycznego klucza szyfrującego.

BIBLIOGRAFIA

- [1] IEEE Standards for LAN and MAN: *Virtual Bridged LAN*. Draft Standard P802.1Q/D9, 30th July 1998.
- [2] Górski G., Woźniak J.: *Analiza metod tworzenia sieci VLAN*. W: Proceedings of VII Międzynarodowa Konferencja Telekomunikacji i Informatyki, October 7-9 1998, CSŁil Zegrze, Poland, t.II, s.65-74.
- [3] Górski G., Woźniak J.: *Dynamic virtual LANs with access authorisation..* W: Proceedings of IX Regional Conference On Military Communication and Information Systems 2000, October 4-6 2000, Zegrze, Poland, t.III s.89-93.
- [4] Górski G., Woźniak J.: *Self-defining virtual LANs (AUTO-VLANs) with access authorisation.* W: Proceedings of the International Scientific NATO PIP/PWP Conference Security and Protection of Information CATE 2001, May 9-11 2001, Brno, Czech Republic., t.I, s.22-28.
- [5] Banel K.: *Zagrożenie bezpieczeństwa sieci w warstwie 2*. NetWorld 11/2003(95), IDG Poland S.A., 2003.
- [6] Adams C., Lloyd S.: *Understanding Public-Key Infrastructure*, Macmillan Technical Publishing, 2001.

DYNAMIC VIRTUAL LANs WITH IMPLEMENTATION OF PUBLIC-KEY INFRASTRUCTURE**Summary**

The paper presents security analysis of data transmission in dynamic virtual LAN with access authorisation. Implementation of secure user access to network resources is possible basing on the described protocol which assumes that chosen cryptographic algorithms are used to secure data transmission between distributed elements of the protocol. First of all Public Key Infrastructure (PKI) was used to improve a level of security. Cipher algorithms derived from X.509 standard processing asymmetric cryptographic keys are especially useful for dynamic VLANs with access authorisation. Presented protocol assumes that before authorisation a user has no information about group session keys used for data encryption and decryption. User login procedure based on PKI first authenticates a new group member and then securely sends him the group session key.

Samer Jamous

Katedra Systemów i Sieci Telekomunikacyjnych, Politechnika Gdańska

**TRANSFER OF CHARGING INFORMATION BETWEEN
EXCHANGES OF DIFFERENT NETWORK OPERATORS
IN MULTI-PROVIDER ENVIRONMENT**

Abstract

This article is a review article. It presented different configuration systems of charging in multi-provider environment, analyzed different options of charging information transfer in such environment, and discussed SS7 signaling standards, which have been developed by ETSI and used for charging applications.

1. INTRODUCTION

The exchange of charge information between Network Operators (NOs) and/or Service Providers (SPs) is one of the main questions nowadays in multi-provider environment.

Interconnect accounting, once needed to only handle agreements between a couple of nationally owned carriers, has grown more complex as more network operators enter the market, offering a growing sophisticated services. Competition and liberalization, as it is known in Europe and elsewhere in the world, is altering interconnect accounting in unpredictable ways. In such environment the charging information transport question is becoming increasingly very important.

Interconnection isn't what is used to be – Simple agreements between two national carriers, whereby carrier A simply pays carrier B for terminating its traffic, unconcerned about where the traffic goes from there. The networks today have been changed in numbers and complexity. The networks today can belong to operators in competition. A call can today include a number of network types. The involved networks for a call can today belong to different operators in competition. Services from network operators and/or service providers can be added either somewhere on the traffic route and/or at any time during the same call.

This article focuses on the exchange of charge information in multi-provider environment and studies the Application Transport Mechanism ATM defined by ITU-T, which is used to transport charging information in such environment.

2. CONFIGURATIONS FOR CHARGING IN MULTI-PROVIDER ENVIRONMENT

There are two configurations for charging in multi-provider environment, centralized and distributed. The model of centralized charging has not been implemented yet and there has been little development done with regards to this method. The existing implementations and standards have to do rather with the distributed method. The current implantations of existing distribution method have some limitations and disadvantages which have been discussed here.

2.1. The centralized charging method

Centralized charging means charging outside the switch points in charging centres common to a number of switch points.

It might be a practical advantage for customers to pay all the services related to a specific call in a single bill. The providers could decide to join their efforts to meet customer's wishes. A solution could be the centralized charging method.

There are two approaches to achieve that:

1. For services added at switch points along the route, the usage related data for charging are transferred to the charging centre by data links from the originating switch points SSPs for the services to the charging centre.

2. The usage data is transferred from the service originating switch points to the switch point for the call access.

In relation to a specific call, it is necessary to link together the different originating call related records. For this purpose, a call reference must be inserted into the records. The call reference in the UMR/CDR must be unique in order to identify unambiguously the records belonging to the same call [1].

Below are configurations of centralized charging system:

In Figure 1 the switch points (within the same NO) have individual data links to the mediation device. In order to join the usage metering data coming to the charging centre via different data links from network and service providers contributing to the same call, a call reference to any transfer of UMR/CDR is needed.

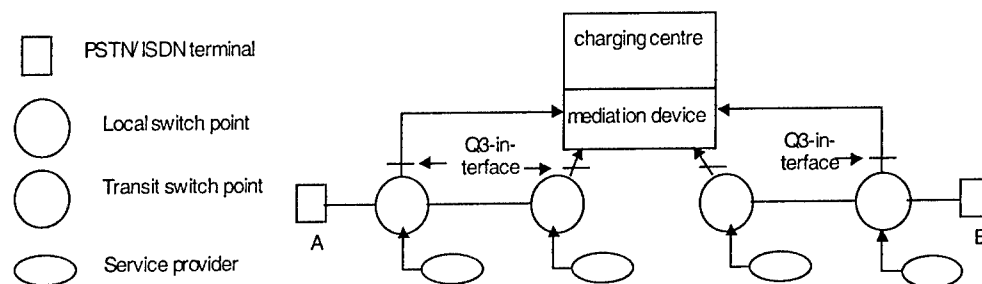


Figure 1: Reference configuration 1 for a centralised charging system

A possible variation is presented in figure 2, where a NO collects data from another NO. outsourcing billing. The major difference between the previous one and this one is in the interfaces for the mediation device: Q3 interfaces (or a proprietary interface) in the network of same organization.

In order to join the usage metering data coming to the charging centre via different data links from network and service providers contributing to the same call, a reference to any transfer of UMR/CDR is needed.

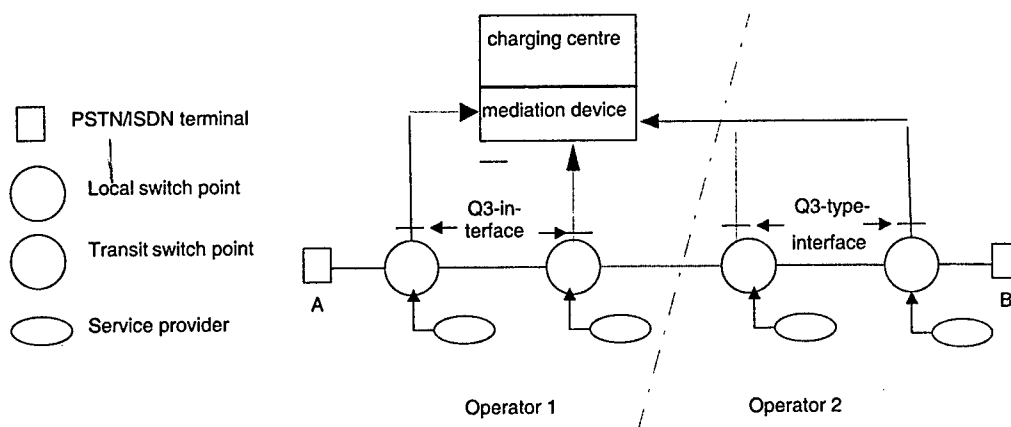


Figure 2: Reference configuration 2 for a centralised charging system

Another possible variation is presented in figure 3, where a NO collects data from another NO via a mediation device.

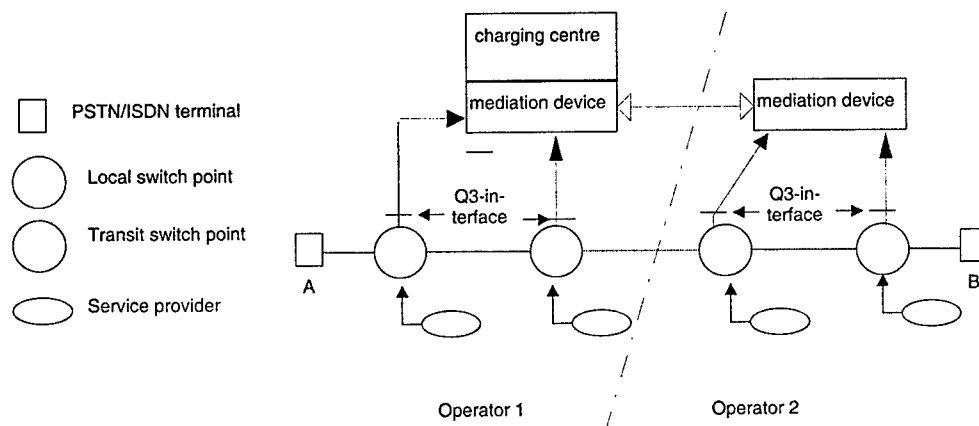


Figure 3: Reference configuration 3 for a centralized charging system

2.2. The distributed charging method

The distributed charging method is the traditional charging system suitable for simple charging scheme. Charging is distributed partly or completely to the switch points or SCPs in the network. The switch points are individually able to monitor the calls originated by their customers and to collect and store the call related data, which will be used in the charging process. In some cases price calculation is managed in the switching point itself, in other cases managed externally based on usage metering records transferred [1].

For each call a charging center and a billing center can be identified. The charging center or the charge determination point is the exchange responsible for determining the charge to be applied, and for applying that charge.

The billing center or the registration point is the exchange responsible for recording the applied charge against a subscriber register. It also applies increments to a home meter or payphone. The billing and charging center can be either in the same exchange or different exchanges.

When the charging center and billing center are in different exchanges, and the charge for the call is to be recorded, a mechanism must exist to transfer the charging information from the charging center to the billing center. This is achieved by either sending meter signals down the trunks (metering trunks), or the information may be passed using a signaling tariff message as used by Common Channel Signaling (CCS) e.g. SS7, ISUP, TUP.

Some operators have extended the traditional line ISUP signalling scheme by proprietary solutions for sending charging information from the charging determination point to the call originating SSP. Signalling systems to handle this have been specified in ES 201 296: ISDN User Part (ISUP) Support of Charging [2]. It specifies a method for transferring tariff information between networks using messages in Signalling System No7. The standard is designed to be used in situation, where the details of the tariff charged to the caller depend on the terminating network. An example is where the originating network charges the call termination charge of the terminating network plus a mark-up. It also covers the situation, where the tariff rate may change during a call. But it may take several years before this standard is implemented widely in networks. In Poland this standard is still not used by the Polish Network Operators.

Complicated scenarios in multi-provider environments could require charging information, which is not detectable at the Charging Center – Charging Determination Point. An example is a call getting service additions during the call from distant network/service providers, or the call can follow different routes with different charging. Complicated scenarios will hardly be possible to manage in the distributed charging method.

In multi-provider environments the distributed charging method would probably only be able to support the providers individually and not to support a charging system for joint use by a group of providers.

The next sections of this article will deal with the implementation of the distributed charging method in terms of signaling.

3. DISTRIBUTED CHARGING AND ISUP VERSIONS ISSUES

ISDN User Part (ISUP) of the Signalling System No. 7, version 1 (ISUP1) has no charging signaling information. The tariff determination takes place in the local exchange or the first exchange in the path of exchanges built with SS7 signaling, based on destination number analysis.

In ISUP2, there is a possibility of sending the charging information via charge message (CRG message) down the trunk from the point, which determines the charge (Charging Centers) to the exchanges, which are down in the network hierarchy. However this charge information has local or country application. Its format has not been defined in ETSI standards. It is different in various countries - the format in Poland is different than Netherlands or India.

Both versions are not good for multi-provider environment – not flexible and can be the source of many disagreements between Network Operators.

In next version of ISDN User Part – ISUP3, an enhancement has been added, which is the Application Transport Mechanism – APM, for use by applications requiring a bearer in conjunction with the support of the application's signaling information flow. APM ensures signaling connection between two different applications implemented in the switches in the network (APM Users) by the use of additional parameter APP (one or more) in ISUP messages, and if necessary in additional APM and PRL messages.

Treating the tariff function and charging as an application (APM user), which needs to be located in selected switches of the network – national and international, ETSI has defined signaling for this application, enabling the realization of multi-provider environment and full realization of the Advice of Charge service.

3.1. Application Transport Mechanism

This mechanism has been defined in ITU-T Q.765 (05/98) [3] and called APM "98"-user application, and in ITU-T Q.775 (06/2000) referred as APM "2000"-user application [4], and after its modification was approved by ETSI as EN standard in EN 301 069-1 (10/99).[5]

The Application Transport Mechanism (APM) is capable of creating signalling associations between APM-user application logic located at a Public Initiating Node (PIN) and its peer APM-user application logic located at the Public Addressed Node (PAN). The ability to create different PIN/PAN relationships is only limited by the addressing mechanism used. At call setup, the ISUP basic call mechanism is used to route the call through the network towards the PAN in the same manner as it normally establishes a call. The APM addressing mechanism is used to determine the node in the call path where the Application Information has to be handled but has no influence on the routing of the call [6].

At any time of the call, it is possible to create the following PIN/PAN relationships:

- PIN and PAN at end exchanges,
- PIN at transit exchange and PAN at end exchange,
- PIN at end exchange and PAN at transit exchange,
- PIN and PAN at transit exchanges.

It means that a PIN located anywhere in the call path has the capability to create a PIN/PAN relationship with any exchange in the call path having the APM-user capability. The intermediate exchanges having the APM capability but not the APM-user capability will transit the APM information.

It is, of course, possible to communicate in both directions once a PIN/PAN relationship has been created. It should be noted that only one PIN/PAN relationship may exist over the same segment of the call path for a given application (Context identity value), which does not use the explicit addressing mechanism. For example, if a PIN/PAN

relationship has been established (PIN) between the Originating Local Exchange (OLE) and an intermediate exchange in the call path (PAN), it is not possible to later establish a PIN/PAN relationship from the destination local exchange to the originating local exchange for the same application, since its corresponding PAN would be located at the OLE and would therefore result in two PIN/PAN relationships overlapping at the call leg between the OLE and the intermediate exchange. However, it is possible to establish overlapping PIN/PAN relationships for a given application using the explicit addressing mechanism [4].

3.2. Charging information transport based on ETSI standards – ES 201 296

The last version of this standard is v1.3.1 (2003-02). The standard specifies the procedures of transfer of charging related information between a charge determination point, a connection control point and a charge registration or charge generation point.

A charge determination, connection control, charge registration and charge generation points of the call may be located within the network of one operator (single network operator environment) or may be located in different networks of different operators (multi-operator environment).

The configuration of several charge determination points for one call is possible. It is assumed that there is only one CGP for the call. Moreover, it is assumed for the present document that there is either one CRP for each CDP or one CRP for all CDPs.

The transferred tariff information represents direct tariff or add-on charge (no pointers to charging data), either in meter-pulse format or in currency format. The interworking with the conversion into charge band numbers or meter-pulse units is not supported.

A charge determination, registration and generation points of the call may be located within the network of one operator (single network operator environment) or may be located in different networks of different operators (multi-operator environment).

The configuration of several charge determination points for one call is possible. It is assumed that there is only one CGP for the call. Moreover, it is assumed for the present document that there is either one CRP for each CDP or one CRP for all CDPs.

The concept is illustrated by the following example.

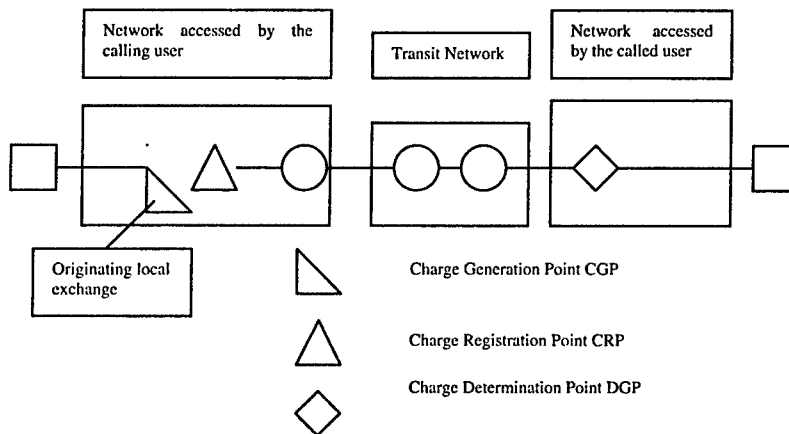


Figure 4: Configuration of CDR, CRP and CGP

3.3. ETSI – ES 201 296 Charging issues

The ETSI charging mechanism adds flexibility to the existing charging mechanism. However, this form for charging has some serious limitations, which make its application very difficult in some environments. They are discussed below.

3.3.1. Implementation issues

The ETSI charging method requires that all the exchanges in the path of the call support SS7 and ETSI charging mechanism. ISUP protocol with APM mechanism should be implemented.

The exchanges, which will determine the charge and register the charging information must have an application, which can work with APM mechanism.

3.3.2. Operational issues

Presently, there seems to be only a few operators in the world who has implemented an ETSI Charging system due to the strict network environment requirements. If an operator were to attempt to do so, it would need that other interconnecting operators implement this system.

3.3.3. Administrative issues

By implementing an ETSI Charging scheme, Access Providers will have to undertake the full charging responsibilities of all the service providers interconnected to their telecom network. It would mean that the Access Provider will bear all the operational and administrative costs and burden of providing the customer care & billing to all end-users, regardless of the fact that the actual amount of its network resources which were used in completing calls may have in fact been minimal. This will further complicate the revenue sharing.

3.3.4. Standardization issues

It should be noted that there are network requirements and signalling limitations that are not covered in the ETSI standard like:

- which currency is used,
- complaint handling between network operators in case of incorrect advice of charge information,
- currency exchange in case of international interconnection.

4. CONCLUSION

In multi-provider environment as it is today, the exchange of charging information between NOs/SPs is becoming increasingly critical. Most of Network Operators don't have implantations for flexible charge determination between themselves. Most of them still use traditional methods of interconnection, which do not fix in multi-provider environment as it is today. There are two configuration systems for multi-provider environment – Distributed Configuration System and Centralized Charging System. There is no group of Network Operators, who uses currently the Centralized Charging Method, but there are some Operators who use the Distributed Charging method. Distributed Charging method, which uses Charge message e.g. CRG message is not a flexible system. Few countries only have it implemented. The Distributed System, which is covered in the ISUP3 protocol using the APM mechanism is better system. This system however has some disadvantages and limitations which have been discussed in this white paper. The centralized Charging System has not been implemented and still need much research and studies.

REFERENCES

- [1] ETSI: TR 101 619 v.1.1.1 (1998-11); *Network Aspects; Consideration on network mechanisms for charging and revenue accounting*, ETSI, 1998.
- [2] ETSI: ES 201 296 v1.3.1 (2003-02); *Integrated Service Digital Network (ISDN); Signalling System No.7 (SS7); Signalling aspects of charging*, ETSI, 2003.
- [3] ITU-T: Q.765 (05/98); *Specification of Signalling System no. 7 – Application transport mechanism*; ITU-T, 1998.
- [4] ITU-T: Q.765 (06/2000); *Specification of Signalling System no. 7 – Application transport mechanism*; ITU-T, 2000.
- [5] ETSI: EN 301 069-1 V1.3.1 (2001-02); *Integrated Service Digital Network (ISDN); Signalling System No.7 (SS7); Application transport mechanism*, ETSI, 2001

PRZEPŁYW DANYCH ZALICZENIOWYCH MIĘDZY RÓŻNYMI OPERATORAMI TELEKOMUNIKACYJNYMI W ŚRODOWISKU WIELOOPERATORSKIM

Streszczenie

Jest to praca przeglądowa. W pracy tej przedstawiono różne konfiguracje systemów taryfikacji i zaliczania w środowisku zdemonopolizowanym. Przeanalizowano różne opcje przepływu danych w takim środowisku i omówiono rozszerzenie sygnalizacji SS7, która jest opracowana przez ETSI i wykorzystywana dla aplikacji zaliczania.

Sylwester Kaczmarek, Marcin Narloch

Katedra Systemów i Sieci Telekomunikacyjnych, Politechnika Gdańska

BADANIE WPŁYWU RUCHU NIŻSZEGO PRIORYTETU NA JAKOŚĆ USŁUG O WYMAGANIACH CZASU RZECZYWISTEGO W SIECI IP QOS

Streszczenie

W referacie przedstawiono wyniki badań dotyczące wpływu rozkładu długości pakietów ruchu o niższym priorytecie na jakość usług o wymaganiach czasu rzeczywistego (*real time services*), korzystających ze schematu obsługi Expedited Forwarding Per Hop Behavior (EF PHB) w sieci IP QoS Differentiated Services (DiffServ). Wyniki zaprezentowane w referacie zostały uzyskane zarówno przy pomocy modeli analitycznych, opisujących gwarancje statystyczne dla ruchu EF (aproksymacja rozkładu opóźnień pakietów strumieni EF), jak i przy pomocy modelu symulacyjnego sieci DiffServ, stworzonego w symulatorze NS2. Wnioski z otrzymanych wyników przedyskutowano w kontekście algorytmów sterowania ruchem w sieci IP QoS DiffServ.

1. WSTĘP

Nowoczesna wielousługowa sieć telekomunikacyjna musi zapewniać realizację różnorodnych usług przy pomocy jednolitej platformy [1]. Usługi w takiej sieci obejmują zarówno przenoszenie danych, których użyteczność jest uwarunkowana przede wszystkim spełnieniem ściśle sprecyzowanych wymagań związanych z czasem przesłania od końca do końca sieci (reżim czasu rzeczywistego), jak i przenoszenie danych, których użyteczność określa jedynie zawartość, bez dokładnego określenia chwili ich dostarczenia. W pierwszym przypadku mamy do czynienia z usługami typowymi dla dotychczasowych sieci telekomunikacyjnych, natomiast w tym drugim przypadku mamy do czynienia z usługą, która jest charakterystyczna dla tradycyjnej sieci IP, w której brak jest jakichkolwiek gwarancji kiedy i czy w ogóle pakiety zostaną dostarczone przez sieć (usługa best-effort).

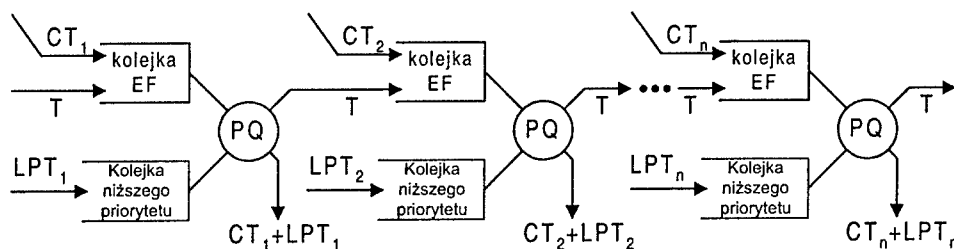
Obecny rozwój sieci telekomunikacyjnych wskazuje na to, że wspólną platformą dostarczającą możliwości realizacji wielu usług, w tym usług z gwarantowaną jakością obsługi będzie sieć IP QoS. Sieć ta będzie wykorzystywać protokół IP wraz ze zbiorem odpowiednich mechanizmów określających modele realizacji usług z gwarantowaną jakością obsługi. Obecnie największą uwagę skupiają możliwości zaproponowanego przez Internet Engineering Task Force modelu IP QoS Differentiated Services (DiffServ) [2,3], a w szczególności problem skutecznej realizacji w sieci DiffServ usług o sprecyzowanych

wymaganiach czasowych (real-time services) wykorzystujących dedykowany dla tego celu schemat obsługi pakietów w węzłach DiffServ nazywany Expedited Forwarding Per Hop Behavior (EF PHB) [4,5,6]. Spośród metod analizy sieci z agregacją i wspólną obsługą w jednej kolejce wielu strumieni (takich jak np. DiffServ), interesująco z praktycznego punktu widzenia przedstawia się podejście stochastyczne [7], które zapewnia wyższy poziom wykorzystania sieci niż podejście deterministyczne [8,9].

Badania przedstawione w referacie stanowią kontynuację i rozszerzenie badań przedstawionych w pracach [10,11,12]. Organizacja referatu jest następująca. W rozdziale drugim przedstawiony został model badanej sieci. Rozdział trzeci zawiera prezentację wyników, natomiast podsumowanie pracy znajduje się w rozdziale czwartym.

2. MODEL BADANEJ SIECI

W celu określenia rozkładu opóźnień pakietów strumieni wykorzystujących schemat obsługi EF PHB w sieci DiffServ została dokonana analiza sieci przedstawionej na rys. 1. Sieć ta składa się z n szeregowo połączonych priorytetowych systemów kolejkowych PQ [1] (węzłów) obsługujących w kolejce o wyższym priorytecie obserwowany strumień T typu CBR przechodzący przez wszystkie węzły oraz interferujący strumień CT_k o poissonowskiej charakterystyce i wielkości pakietów MTU przechodzący przez k -ty węzeł. Wykorzystywane jest tu założenie o „pomijalnym jitterze” [7,13] w opisie usług EF oraz wnioski płynące z analizy parametrów charakteryzujących ruch w sieci IP uzyskanych na drodze pomiarów rzeczywistych sieci [14,15]. W kolejce o niższym priorytecie obsługiwany jest ruch nie mający wymagań czasu rzeczywistego, korzystający z innych PHB niż EF PHB [3]. Z uwagi na przyjęcie systemu PQ jako sposobu realizacji EF PHB, z punktu widzenia obsługi ruchu o wyższym priorytecie kolejka niższego priorytetu stanowi reprezentację zarówno pojedynczego bufora pakietów, jak i złożonego systemu kolejkowego, np. WFQ [1].



Rys.1. Sieć priorytetowych systemów kolejkowych połączonych szeregowo.

W modelu badanej sieci wpływ ruchu o niższym priorytecie na jakość obsługi dla strumieni EF jest uwzględniany w postaci modelu kolejkowego systemu z wielokrotnymi wakacjami. W systemie z wakacjami proces przybyć i obsługi ma taką samą charakterystykę jak w klasycznym systemie kolejkowym. Jednakże, w modelu z wakacjami, jeżeli kolejka o wyższym priorytecie (kolejka EF) jest pusta to serwer udaje się na wakację. Czas trwania wakacji w analizowanym modelu jest równy czasowi transmisji pakietu niższego priorytetu, tak więc w zależności od rozkładu długości pakietów niższego priorytetu mamy w ogólności do czynienia z wakacjami o różnym czasie trwania. Jeżeli kolejka o wyższym priorytecie jest pusta po powrocie serwera z wakacji, to serwer ponownie udaje się na wakację (wielokrotne wakacje). Gdy serwer systemu kolejkowego po powrocie z wakacji

znajdzie w kolejce EF oczekujące pakiety, obsługuje je do momentu opróżnienia kolejki EF (dyscyplina wyczerpująca). Taki system kolejkowy pozwala na modelowanie rzeczywistego systemu z kolejką priorytetową i łączem przenoszącym pakiety nie-EF w każdej chwili, gdy nie ma pakietów EF do przeniesienia. Podejście to jest oparte na zasadzie najgorszego przypadku w odniesieniu do jakości strumienia EF, ponieważ zakłada się, że w węźle jest zawsze pakiet niższego priorytetu do wysłania.

W referacie do określenia rozkładu opóźnień pakietów wykorzystywana jest metoda oparta na aproksymacji gaussowskiej [10,16] oraz metody oparte na Teorii Wielkich Odchyłeń [17], tj. aproksymacja wykorzystująca twierdzenie Czernowa i aproksymacja wykorzystująca twierdzenie Bahadur-Rao. Powyższe aproksymacje zostały dokładnie omówione w pracy [11]. W każdej z tych metod niezwykle istotną rolę odgrywa znajomość funkcji tworzącej momenty rozkładu opóźnienia dla danego modelu kolejkowego.

Funkcja tworząca dla kolejki z wakacjami może zostać określona dzięki własności stochastycznej dekompozycji opisanej w [18,19] i uogólnionej w [20,21]. Własność stochastycznej dekompozycji pozwala traktować czas oczekiwania w kolejce M/GI/1 z wakacjami jako sumę dwóch niezależnych składników, jednego o rozkładzie czasu oczekiwania w zwykłej kolejce M/GI/1 bez wakacji, a drugiego będącego resztkowym czasem trwania wakacji. Tak więc, funkcja tworząca dla systemu M/D_{MTU}/1 z wakacjami może zostać określona jako:

$$M(\theta) = \frac{V(\theta) - 1}{v \cdot \theta} M_{M/D/1}(\theta), \quad (2.1)$$

gdzie $M_{M/D/1}(\theta)$ jest funkcją tworzącą w zwykłej kolejce (bez wakacji), $V(\theta)$ oznacza funkcję tworzącą rozkładu czasów wakacji, zaś v średni czas trwania wakacji. W badanym systemie przyjęto dyskretny rozkład prawdopodobieństwa wystąpienia kilku określonych długości pakietów niższego priorytetu, a więc funkcja tworząca rozkładu czasów wakacji przyjmuje postać:

$$V(\theta) = \sum_{k=1}^K v p_k \cdot \exp(\theta \cdot v_k), \quad (2.2)$$

gdzie v_k jest czasem trwania wakacji równym czasowi transmisji pakietu k-tego typu (k-tej długości), natomiast $v p_k$ jest prawdopodobieństwem wystąpienia pakietu k-tego typu w strumieniu niższego priorytetu.

W modelu sieci zostało założone, że natężenie ruchu oferowanego ρ_T obserwowanego strumienia jest stosunkowo niewielkie w odniesieniu do natężenia ruchu ρ_{CT} w każdym węźle. Ponadto zakłada się, że kolejki w poszczególnych węzłach są niezależne (czas oczekiwania pakietów EF w każdej z kolejek jest niezależną zmienną losową).

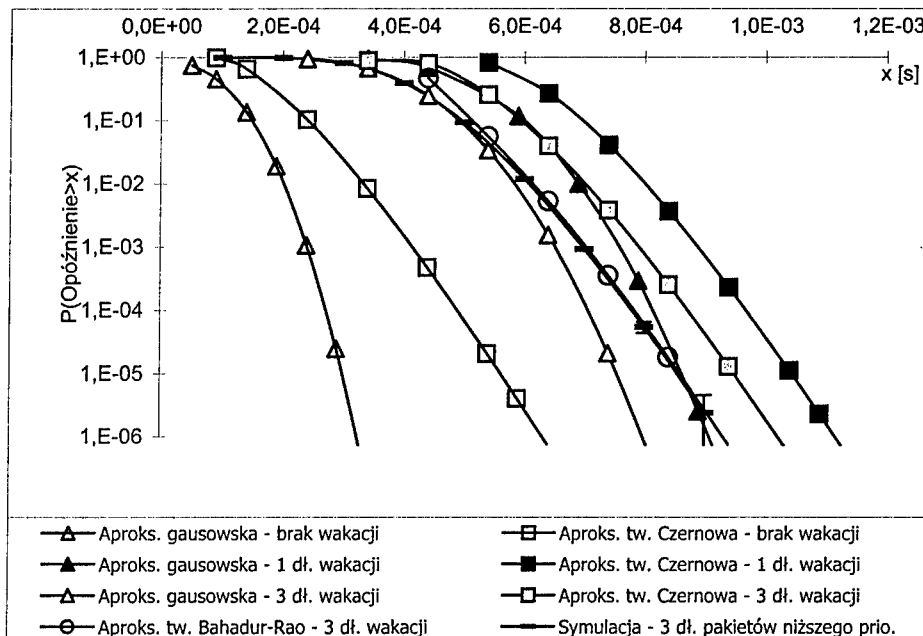
3. WYNIKI NUMERYCZNE

W celu weryfikacji dokładności aproksymacji rozkładu opóźnień pakietów uzyskiwanej przy pomocy stworzonego modelu analitycznego zostało dokonane porównanie wyników teoretycznych z wynikami eksperymentów symulacyjnych. Przeanalizowano sieć składającą się z dziesięciu węzłów z łączami międzywęzłowymi o przepływności 150 Mb/s. Jako ruch tłowy CT_i został wykorzystany strumień Poissona pakietów o wielkości MTU 1500 bajtów i natężeniu ruchu oferowanego ρ_{CT} równym odpowiednio 0,1, 0,3 oraz 0,5.

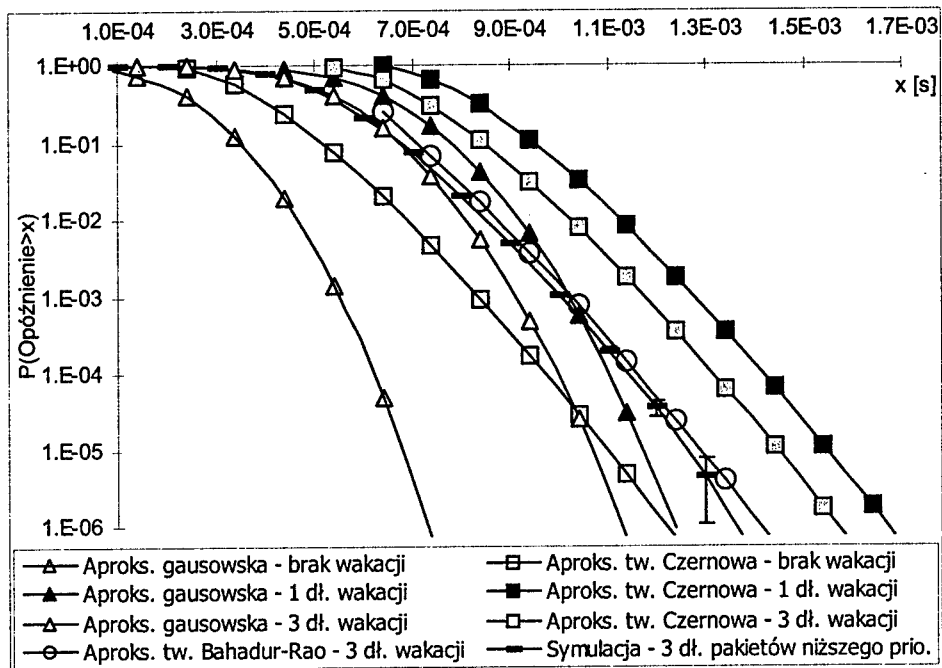
Ruch obserwowany składał się z jednego strumienia CBR o przepływności 1,5 Mb/s (natężenie ruchu oferowanego ρ_T równe 0,01). Ruch o niższym priorytecie składał się z trzech typów pakietów o długościach: 50, 600 lub 1500 bajtów, których prawdopodobieństwo wystąpienia w strumieniu ruchu o niższym priorytecie wynosiło odpowiednio: 0,5, 0,3, 0,2. Należy zauważyć, że długości pakietów oraz rozkład prawdopodobieństwa ich wystąpienia są zgodne z wynikami najnowszych pomiarów sieci IP, przedstawionymi w pracach [23-26].

Rozkład prawdopodobieństwa opóźnień pakietów ruchu obserwowanego w badanej sieci dla odpowiednich wartości ρ_{CT} przedstawiono na rys. 2, 3, 4. Ponadto na rys. 5 przedstawiono rozkład prawdopodobieństwa opóźnień pakietów ruchu obserwowanego w sieci pięciu węzłów z łączami międzywęzłowymi o przepływności 150 Mb/s, w której obciążenia ruchem tłowym ρ_{CT} są niejednorodne dla poszczególnych węzłów i wynoszą $\rho_{CT,i} = i \cdot 0,1; i=1 \dots 5$. Wyniki uzyskane dla przypadku trzech długości pakietów niższego priorytetu oznaczone są jako „- 3 dł. wakacji”.

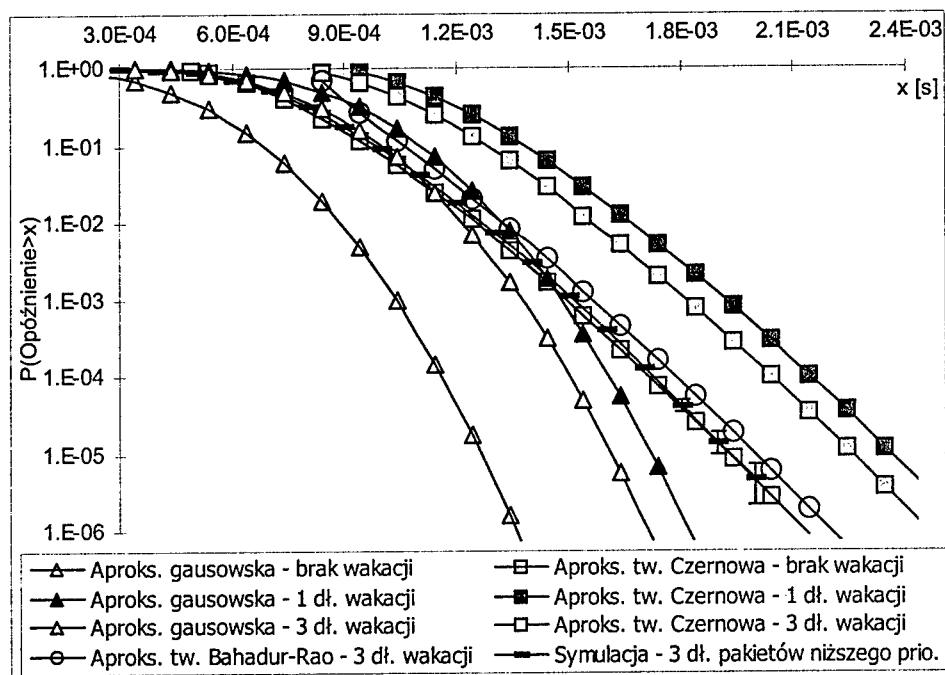
Na rys. 2-5 zostały także przedstawione wyniki badań symulacyjnych wraz z przedziałami ufności obliczonymi dla poziomu ufności równego 0,95. Wyniki te oznaczone są na rys. 2-5 jako „Symulacja - 3 dł. pakietów niższego priorytetu”. W celu umożliwienia porównania otrzymanych wyników na rys. 2-5 przedstawiono dodatkowo aproksymacje rozkładów opóźnień pakietów wynikające z dwóch skrajnych podejść do wpływu ruchu o niższym priorytecie na jakość usług o wymaganiach czasu rzeczywistego. W pierwszym z tych podejść optymistycznie zakłada się model sieci, w którym nie uwzględniono wpływu ruchu o niższym priorytecie na jakość dla strumieni korzystających z EF PHB (systemy kolejkowe $M/D_{MTU}/1$ bez wakacji). Na rys. 2-5 wyniki uzyskane dla takiego założenia oznaczone są jako „- brak wakacji”.



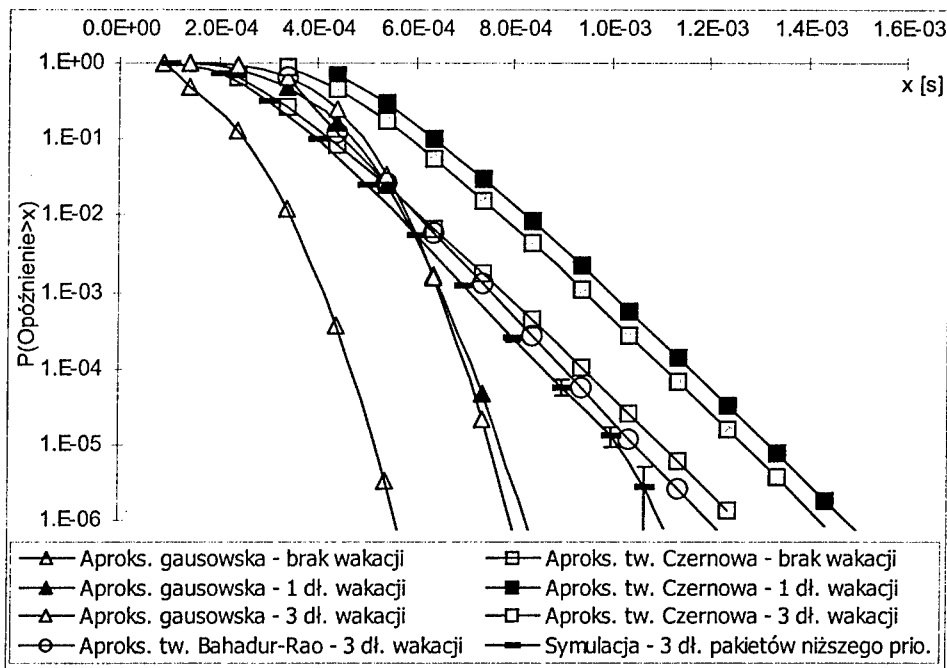
Rys.2. Rozkład opóźnień pakietów strumienia CBR ($\rho_{CT}=0,1; n=10$).



Rys.3. Rozkład opóźnień pakietów strumienia CBR ($\rho_{CT}=0,3$; $n=10$).



Rys.4. Rozkład opóźnień pakietów strumienia CBR ($\rho_{CT}=0,5$; $n=10$).



Rys.5. Rozkład opóźnień pakietów strumienia CBR ($\rho_{CT,i} = i-0,1; i=1...5$).

W drugim, pesymistycznym podejściu zakłada się długość pakietów niższego priorytetu jako wielkość stałą i równą MTU=1500 bajtów, co oznacza konieczność analizy systemu $M/D_{MTU}/1$ z wakacjami o stałym czasie trwania równym czasowi transmisji pakietu MTU. Na Rys. 2-5 wyniki uzyskane dla takiego podejścia oznaczone są jako „1 dł. wakacji”.

Na rys. 2-5 widoczne jest, że w sytuacji sieci w której ruch o niższym priorytecie (nie-EF) składa się z pakietów o różnych długościach teoretyczny model uwzględniający tę sytuację zapewnia najlepszą aproksymację rozkładu opóźnień pakietów usług korzystających z EF PHB (ruch o wyższym priorytecie). Należy jednakże zauważyć, że w badanej sieci różnica pomiędzy dokładnością aproksymacji zapewnianą poprzez przyjęcie szczegółowego modelu teoretycznego, a aproksymacją w którym wpływ ruchu o niższym priorytecie jest modelowany na zasadzie najgorszego przypadku poprzez założenie stałej wielkości pakietów niższego priorytetu jest stosunkowo nieduża. Ponadto, różnica ta staje się tym mniejsza im większe jest obciążenie węzła ruchem wyższego priorytetu (ruch EF). Z drugiej strony pominięcie wpływu ruchu o niższym priorytecie może prowadzić do niedokładnych wyników i nadmiernie optymistycznych oszacowań rozkładu opóźnień pakietów usług o reżimie czasu rzeczywistego korzystających z EF PHB w sieci IP QoS.

Należy zauważyć, że przyjęcie dokładnego modelu rozkładu wielkości pakietów oznacza konieczność znaczącego skomplikowania modelu teoretycznego aproksymacji rozkładu opóźnień pakietów EF. Ponadto, taki model wymaga dokonania określonych założeń odnośnie parametrów rozkładu wielkości pakietów. Na obecnym etapie rozwoju sieci IP podstawowym źródłem informacji o parametrach tego rozkładu są pomiary ruchowe dokonywane w rzeczywistych sieciach. Ponieważ jednak dokładne wyniki tych pomiarów dostępne są *post factum*, a biorąc pod uwagę niezwykle dynamiczny, rewolucyjny zakres zmian w sieciach IP i aplikacji z niej korzystających, wyniki te obarczone są

podstawową wadą – opisują sieć, której bieżące charakterystyki są znacząco inne od opisu uzyskanego z pomiarów. Z praktycznego punktu widzenia oznacza to, że wpływ ruchu o niższym priorytecie na jakość usług EF musi być uwzględniany, ale wystarczający stopień dokładności zapewnia przyjęcie w modelach teoretycznych najgorszego przypadku wielkości pakietów niższego priorytetu, czyli wielkości MTU.

4. PODSUMOWANIE

W zakresie realizacji usług o wymaganiach czasu rzeczywistego wpływ ruchu o niższym priorytecie nie może zostać pomijany. W związku z tym wymagany jest opis tego oddziaływania i powiązanie go z metodami opisu jakości tych usług [27,28,29]. Na podstawie przedstawionych w referacie wyników satysfakcjonującym, z punktu widzenia dokładności i złożoności, podejściem do oceny wpływu ruchu niższego priorytetu jest podejście oparte na założeniu najgorszego przypadku wielkości pakietów niższego priorytetu. Podejście to upraszcza problem określania rozkładu opóźnienia pakietów usług czasu rzeczywistego w sieci IP QoS DiffServ ze schematem obsługi EF PHB i pozwala na wykorzystanie tej miary oraz miar pochodnych [12] w podstawowych z punktu widzenia gwarancji jakości obszarach zastosowań takich jak: sterowanie przyjęciem zgłoszeń, ruting, wymiarowanie i projektowanie sieci.

BIBLIOGRAFIA

- [1] Roberts J. et al. (ed.), *Broadband Network Teletraffic. Performance Evaluation and Design of Broadband Multiservice Networks*. (Final Report of COST 242), Springer Verlag, Heidelberg, October 1996.
- [2] Blake S. et al., *An Architecture for Differentiated Services*, RFC 2475, December 1998.
- [3] Nichols K. et al., *A Two-bit Differentiated Services Architecture for the Internet*, RFC 2638, July 1999.
- [4] Davie B. (ed.), et al., *An Expedited Forwarding PHB (Per-Hop Behavior)*, RFC 3246, March 2002.
- [5] Charny A. et al., *Supplemental Information for the New Definition of the EF PHB (Expedited Forwarding Per-Hop Behavior)*, RFC 3247, March 2002.
- [6] Jacobson V. et al., "An Expedited Forwarding PHB", RFC 2598, June 1999.
- [7] Bonald T., Proutiere A., Roberts J. W., *Statistical Performance Guarantees for Streaming Flows using Expedited Forwarding*, Proceedings IEEE Infocom 2001, Anchorage, USA, 2001.
- [8] Bennett J., et al., *Delay Jitter Bounds and Packet Scale Rate Guarantee for Expedited Forwarding*, Proceedings IEEE Infocom 2001, Anchorage, USA, 2001.
- [9] Charny A., Le Boudec J.-Y., *Delay Bounds in a Network with Aggregate Scheduling*, Proc. of QoSIS'2000, Berlin, Germany, September 25-26, 2000.
- [10] Narloch N., Kaczmarek S., *Methods for Evaluation Packet Delay Distribution of Flows Using Expedited Forwarding PHB*, Proceeding of 2nd Polish-German Teletraffic Symposium PGTS2002, Gdansk, September 2002, pp. 85-95.
- [11] Narloch M., S. Kaczmarek, *Problem Jakości Usług w Sieci Opartej na Platformie IP*, Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej, Technologie Informacyjne, Nr 1, Gdańsk 2003, ss. 301-314.
- [12] Narloch N., Kaczmarek S., *Admission Control Method Based on Effective Delay for Flows Using EF PHB*, Proceedings of Workshop on Architectures for Quality of Service in the Internet. Jointly held with the Final AQUILA IST Seminar Art-QoS 2003, Warsaw March 24-25 2003, pp. 125-136. Także w W. Burakowski, B. Koch and A. Beben (eds.): *Architectures for Quality of Service in the Internet*, Springer-Verlag, LNCS Series, volume No 2698, 2003, pp. 127-138.

- [13] Brichet F., Massoulié L., Roberts J., *Stochastic ordering and the notion of negligible CDV*, Proceedings of ITC 15, Elsevier, 1997.
- [14] Cao J., Cleveland W.S., Lin D., Sun D. X., *Internet Traffic Tends Toward Poisson and Independent as the Load Increases*, in Nonlinear Estimation and Classification, eds. C. Holmes, D. Denison, M. Hansen, B. Yu, and B. Mallick, Springer, New York, 2002.
- [15] Cao J., Ramanan K., *A Poisson Limit for the Unfinished Work of Superposed Point Processes*, in Proceedings IEEE Infocom 2002, Tel-Aviv 2000.
- [16] Grossglauser M., Keshav S., *On CBR Service*, Proceedings IEEE Infocom 1996, San Francisco, USA, March 1996 (extended version).
- [17] Dembo A., Zeitouni O., *Large Deviations Techniques and Applications*, Jones and Bartlett Publishers, Boston 1992.
- [18] Fuhrmann S. W., *A Note on The M/G/1 Queue with Server Vacation*, Operations Research, Vol. 32, pp. 1368-1373.
- [19] Fuhrmann S. W., Cooper R. B., *Stochastic Decompositions in a M/G/1 Queue with General-ized Vacation*, Operations Research. Vol. 33, No.5, pp. 1117-1129, 1985.
- [20] Doshi B.T. *A Note on Stochastic Decomposition in a GI/G/1 Queue with Vacations or Set-up Times*, Journal of Applied. Probability Vol. 22, 1985, pp. 419-428.
- [21] Doshi B.T., *Queueing Systems with Vacations - A Survey*, Queueing Systems Theory and Applications, Vol. 1, pp. 29-66, 1986.
- [22] Doshi B.T., *Generalizations of the stochastic decomposition results for single server queues with vacations*, Comm. Statistic. Stochastic Models Vol. 6, No. 2, pp.307-333, 1990.
- [23] Shannon C., Moore D., Claffy K., *Beyond Folklore: Observations on Fragmented Traffic*, IEEE/ACM Transactions on Networking, Vol.10, No.6, December 2002, pp. 709-720.
- [24] Brownlee N., Claffy K., *Understanding Internet traffic streams: Dragonflies and tortoises*. IEEE Communications Magazine, Vol. 40, No. 10, October 2002, pp. 110-117.
- [25] McCreary S., Claffy K., *Trends in Wide Area IP Traffic Patterns: A View from Ames Internet Exchange*, ITC Specialist Seminar on IP Traffic Measurement, Modeling, and Management, Monterey, California, USA, September 14, 2000.
- [26] Claffy K., Miller G. Thompson K., *The Nature of the Beast: Recent Traffic Measurements From An Internet Backbone*, Proceedings of INET98, 1998.
- [27] Karam M., Tobagi F., *Analysis of the Delay and Jitter of Voice Traffic Over the Internet*, Proceeding of Infocom 2001, Anchorage, USA, 2001.
- [28] Goderis D. et al., *Service Level Specification Semantics. Parameters and negotiation requirements*, draft-tequila-sls-01.txt, work in progress, June, 2001.
- [29] Salsano S. et al., *Definition and Usage of SLS in the AQUILA Consortium*, draft-salsano-aquila-sls-00.txt, work in progress, November 2000.

EVALUATION OF LOWER PRIORITY TRAFFIC INFLUENCE ON PERFORMANCE GUARANTEES OF REAL-TIME SERVICES IN IP QOS NETWORK

Summary

In the paper results of performance guarantees evaluation for real-time services using Expedited Forwarding Per Hop Behavior (EF PHB) in IP QoS Differentiated Services (DiffServ) network are presented. Evaluation regards the influence of lower priority packet size distribution on quality of service issues of EF PHB flows. Presented results were obtained with aid of analytical models describing statistical performance guarantees for EF streams (approximation of EF streams packet delay distribution), as well as simulation model of Differentiated Services network developed with aid of NS2 simulation environment. Results of evaluation are discussed in the context of their consequences for traffic control algorithms for IP QoS DiffServ network.

Sylwester Kaczmarek¹, Krzysztof Nowak²

¹) Systemy i Sieci Telekomunikacyjne, Politechnika Gdańska, ²) Siemens S.A., Lizbona

MODEL SYMULACYJNY I BADANIA SIECI MPLS

Streszczenie

Multiprotocol Label Switching (MPLS) to obiecująca technika, umożliwiająca wsparcie jakości usług w sieciach szkieletowych opartych na protokole IP. Aby skorzystać z jej potencjalnych zalet, należy opracować skuteczne procedury i metody gwarantujące realizację jakości usług. Jednym z kluczowych zagadnień w tym obszarze badań jest kierowanie ruchem (ruting), a niezbędnym narzędziem odpowiedni program symulacyjny do modelowania zjawisk w sieci MPLS. W artykule mówiono implementację takiego programu i wyniki uzyskane za jego pomocą. Zwrócono uwagę na specyficzne cechy sieci MPLS i ich wpływ na sposób przeprowadzania symulacji. Dokonano oceny kilku prostych metod kierowania ruchem wykorzystanych w programie. Przedstawiono także kierunki dalszych prac w tym obszarze badań.

1. TECHNIKA MPLS

Technika MPLS została opracowana w połowie lat 90-tych jako rozwinięcie kilku konkurencyjnych technik komutacji etykietowej dla sieci IP. Powodem poszukiwania nowych rozwiązań dla dobrze rozwijającej się platformy IP były problemy z rutingiem. W obliczu bezprecedensowego rozwoju sieci Internet istniejące urządzenia zaczęły mieć problemy wydajnościowe. W tej sytuacji rozwiązaniem miało być zastosowanie komutacji etykietowej. Jednak w trakcie zaawansowanych prac nad nową techniką zaszły wielkie zmiany w technologii procesorów i specjalizowanych układów scalonych. Ich zastosowanie pozwoliło na rozwiązanie wielu problemów, w tym kierowania pakietów. Mimo tego nie zaprzestano prac nad MPLS, gdyż pojawiło się zapotrzebowanie na efektywne zarządzanie ruchem [1], w czym technika ta mogła być niezwykle użyteczna. Duże znaczenie dla kontynuowania prac miała też łatwość budowania sieci wydzielonych [2].

W technice MPLS zwykłe pakiety IP są opatrzone dodatkowym nagłówkiem, tzw. etykietą. Pojedyncze strumienie IP są grupowane w ścieżki LSP (*Label Switching Path*), a wszystkie pakiety należące do tej samej ścieżki mają jednakową etykietę. W przeciwieństwie do protokołu IP, gdzie w czasie transportu informacja adresowa pozostaje niezmienną, MPLS opiera się na komutacji etykiet. Podobnie jak identyfikatory VCI/VPI w ATM, etykiety podlegają zmianie w każdym węźle na drodze połączeniowej.

Etykieta identyfikuje jednoznacznie klasę FEC (*Forwarding Equivalence Class*). Wszystkie pakiety należące do danej klasy FEC są kierowane do tego samego prze-

znaczenia i są traktowane w ten sam sposób, a w szczególności należą do jednej klasy ruchu. Poza kilkoma specyficznymi przypadkami ruting odbywa się wyłącznie w oparciu o etykiety, a nie nagłówki IP.

Budowanie tuneli i sieci wirtualnych jest ułatwione dzięki mechanizmowi stosu etykiet. Polega on na dodawaniu kolejnych etykiet do już zaetykietowanego pakietu. Nowo dodana etykieta „przykrywa” pozostałe i aż do jej usunięcia tylko ona bierze udział w rutingu. Liczba etykiet na stosie jest teoretycznie nieograniczona.

2. MODEL SYMULACYJNY

Zadaniem programu jest odtworzenie i symulacja sieci MPLS oraz wykonywanie pomiarów jakości uzyskiwanej dla założonej charakterystyki ruchu i topologii. Przy jego opracowywaniu należało zaimplementować pracę komutatora etykiet i źródeł ruchu IP oraz połączyć obiekty logicznymi łączami. Wszystkie obiekty muszą być w łatwy sposób konfigurowane, a topologia sieci dowolnie określona. Wymagane jest zbieranie wyników pomiarów parametrów określających jakość i poziom ufności.

Program został napisany w języku C++ z wykorzystaniem mechanizmu rozbudowanej hierarchii klas. Odzworowaniem klas w model sieci są obiekty, które komunikują się ze sobą za pomocą zdarzeń synchronizowanych z wirtualnym czasem symulacji.

Ze względu na prostotę implementacji program nie posiada interaktywnego interfejsu użytkownika. Jest uruchamiany z linii poleceń jako program konsolowy. Wszystkie informacje potrzebne do wykonania symulacji są umieszczone w pliku konfiguracyjnym i w linii poleceń (tylko niezbędne parametry). W trakcie symulacji podstawowe informacje o jej przebiegu kierowane są na ekran. Po jej zakończeniu wyniki są wpisywane do szeregu raportów w postaci plików tekstowych. Ich liczba i zawartość są konfigurowane. Tworzony jest również specjalny plik dziennika (log), gdzie umieszcza się szczegółowe informacje o przebiegu symulacji i ewentualne błędy. Zakres standardowych informacji kierowanych do pliku dziennika jest określany z linii poleceń, natomiast błędy są rejestrowane zawsze.

Źródło generuje pakiety IP o określonej średniej i maksymalnej długości z częstością określoną przez maksymalną¹ i średnią szybkość. W zależności od wartości tych parametrów można modelować źródło CBR (*Constant Bit Rate*) lub VBR (*Variable Bit Rate*). W przypadku źródeł VBR długość pakietu i częstość generacji jest wybierana losowo z rozkładem wykładniczym, przy czym obowiązuje określona maksymalna długość pakietu. Jeśli praca źródła ma być synchronizowana z szybkością łącza, wówczas automatycznie tworzony jest multiplexer, który kolejkuje pakiety.

Model węzła składa się z komutatora i portów wyjściowych. Pakiety z wejść trafiają na komutator, który dokonuje zmiany etykiety i kieruje pakiet do odpowiedniego portu wyjściowego. Komutacja odbywa się w oparciu o szereg tablic zaimplementowanych zgodnie z modelem opisanym w [3]. Jeśli przybyły pakiet nie posiada jeszcze etykiety, wówczas zostaje zakwalifikowany do klasy FEC na podstawie adresu docelowego i klasy ruchu. Następnie na podstawie informacji uzyskanych z tablicy NHLFE (*Next Hop Label Forwarding Entry*) przydziela się pakietowi etykietę i przesyła do odpowiedniego portu. W przypadku pakietów już posiadających etykietę, jej wartość oraz numer wejścia służą jako klucz do tablicy ILM (*Incoming Label Map*). Odczytany zostaje numer wpisu w tablicy

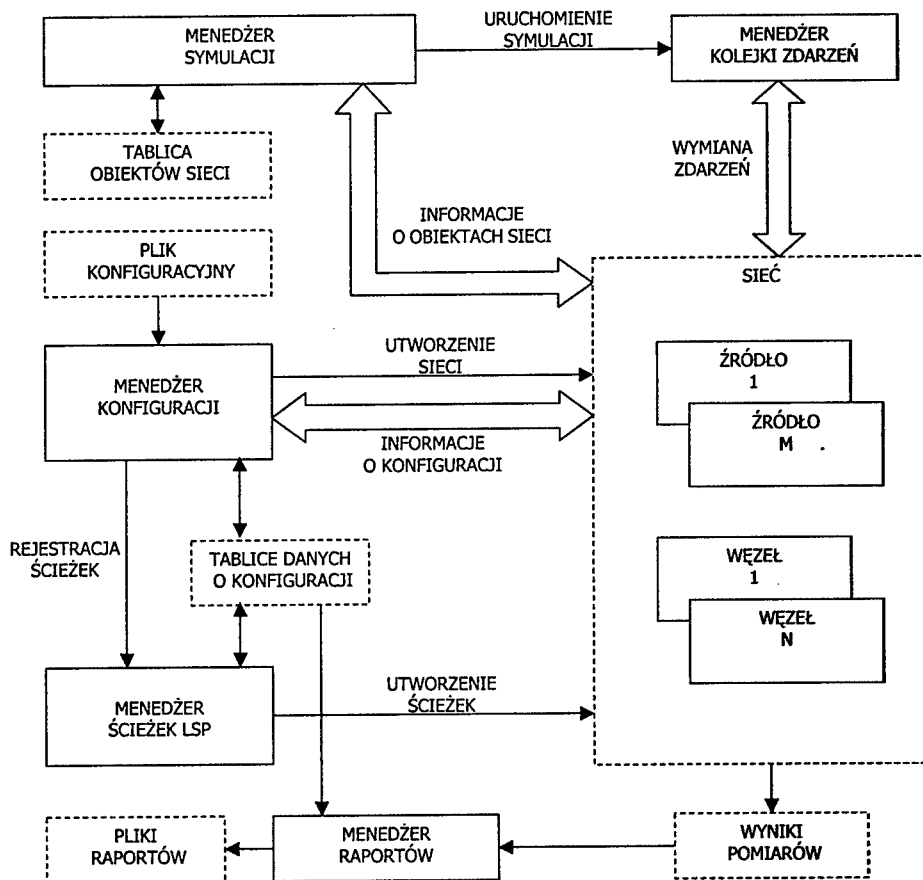
¹ Prędkość maksymalną interpretuje się tu w odniesieniu do średniej długości pakietu. Statystycznie w krótkich odcinkach czasu rzeczywiste pasmo źródła może być większe.

NHLFE, skąd otrzymuje się nową etykietę i port wyjściowy. Po tym pakiet zostaje przesłany do odpowiedniego portu.

Każdy port wyjściowy zawiera zestaw buforów dla każdej klasy ruchu, przy czym klasy dzieli się zgodnie z modelem DiffServ [4]. Klasa EF jest obsługiwana w pierwszej kolejności, aż do opróżnienia kolejki. W drugiej kolejności są obsługiwane pakiety klasy z grupy AF i na końcu pakiety klasy best effort. W ramach klasy AF wybór odbywa się przy pomocy mechanizmu WFQ.

Łącza charakteryzowane są przez przepustowość i czas propagacji. Na opóźnienie pomiędzy momentem opuszczenia przez pakiet komutatora a dotarciem do kolejnego węzła składa się czas oczekiwania w buforze, czas transmitowania pakietu w łącze i czas propagacji.

Oprócz obiektów biorących udział w symulacji istnieje kilka obiektów sterujących, bez których symulacja nie mogłaby działać. Schemat blokowy symulatora przedstawiono na rysunku 1. Poniżej omówiono zadania poszczególnych obiektów.



Rys.1. Obiekty biorące udział w symulacji.

Menedżer symulacji jest najważniejszym obiektem programu symulacyjnego. Zaraz po utworzeniu przejmuje kontrolę nad dalszym przebiegiem programu. Odpowiada m.in. za tworzenie podstawowych obiektów i kontrolowanie czasu symulacji. Przechowuje listę wszystkich obiektów symulacji i ich lokalizację w pamięci. Umożliwia również wymianę danych pomiędzy obiektami, które nie mają ze sobą bezpośrednich powiązań.

Menedżer kolejki zdarzeń zarządza sortowaną listą-kolejką, w której przechowywane są zdarzenia generowane przez wszystkie obiekty. Kluczem do sortowania jest określany przez nadawcę zdarzenia czas dostarczenia. Mechanizm ten zapewnia niezbędną synchronizację pracy niezależnych od siebie obiektów. Symulacja odbywa się tak długo, dopóki w kolejce znajdują się zdarzenia. W normalnej sytuacji nie zdarza się, by kolejka opróżniła się samoistnie z braku nowych zdarzeń. Dlatego aby wymusić zakończenie symulacji, menedżer symulacji nakazuje usunięcie wszystkich zdarzeń z kolejki.

Menedżer konfiguracji odpowiada za odczyt pliku konfiguracyjnego oraz utworzenie źródeł i węzłów. Wszystkie parametry zapisane w pliku są wpisywane do tablic. Dzięki temu plik konfiguracyjny jest odczytywany tylko raz i nie ma potrzeby wykonywać wielokrotnie konwersji z postaci tekstowej, wygodnej dla użytkownika, na binarną, odpowiednią do interpretacji przez obiekty programu. Daje to również możliwość sprawdzenia spójności zadanej konfiguracji. Taka kontrola jest wykonywana tylko raz i pozwala wykryć błędy jeszcze przed rozpoczęciem symulacji.

Menedżer ścieżek LSP zajmuje się tworzeniem ścieżek, na co składają się: określenie drogi, rezerwacja pasma i tworzenie wpisów w tablicach FEC, ILM i NHLFE we wszystkich węzłach na drodze połączeniowej. Ścieżki tworzy się w kolejności numerów przypisanych w konfiguracji.

Menedżer raportów odpowiada za przechowywanie danych pomiarowych i wpisanie ich w odpowiedniej formie do plików tekstowych, które stanowią rezultat pracy programu. W konfiguracji określa się, które z wyników pomiarów są wpisywane do jakich plików.

3. METODY KIEROWANIA RUCHEM

Kierowanie ruchem w sieci MPLS odbywa się na poziomie ścieżek. Nie ma potrzeby podejmowania decyzji o kierowaniu poszczególnych pakietów. Jeszcze przed rozpoczęciem komunikacji następuje utworzenie wszystkich ścieżek, którego najważniejszym etapem jest określenie drogi połączeniowej od węzła początkowego do końcowego, a więc klasyczny ruting sieciowy. Jest to mechanizm statyczny, zdecentralizowany, tzn. bez centralnej tablicy połączeń. Nie ma też centralnego bloku kierowania ruchem a proces wyszukiwania drogi inicjowany jest z węzła wejściowego. Poprzez odpytywanie węzłów dołączonych do kolejnych łączy wyjściowych szuka się wszystkich możliwych dróg połączeniowych. Zbiera się przy tym informację o dostępnym paśmie, wykorzystywaną do wyboru drogi.

Listę dróg sortuje się na bieżąco według zadanego kryterium i wybiera pierwszą z wystarczającym pasmem. Po tym następuje rezerwacja wymaganego pasma na kolejnych łączach i wykreowanie wpisów w tablicach kierowania ruchem NHLFE, ILM i FTN, we wszystkich węzłach na wybranej drodze.

Źródła ruchu przed rozpoczęciem generowania pakietów przydziela się do tak przygotowanych ścieżek. Jeśli odpowiednia ścieżka nie istnieje lub nie ma wystarczającego pasma, wtedy źródło zostaje wyłączone. Pasma wymagane przez źródła VBR wyliczane jest według koncepcji pasma efektywnego metodą aproksymacji gaussowskiej [5].

Wyboru metody kierowania ruchem dokonuje się poprzez określenie kryterium porównywania dróg. Aktualnie dostępne są trzy metody.

1. *Priorytet najkrótszej drogi.* Droga A jest lepsza od B jeśli A przechodzi przez mniejszą liczbę węzłów. Pasma nie jest brane pod uwagę.

2. *Najkrótsza droga z uwzględnieniem dostępnego pasma.* Droga A jest lepsza od B jeśli przechodzi przez mniejszą liczbę węzłów lub przechodzi przez tę samą liczbę węzłów, ale dysponuje większym pasmem.

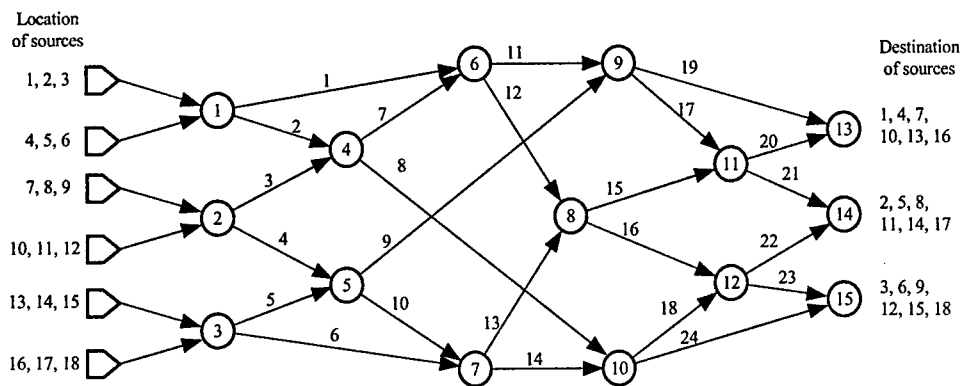
3. *Dostępne pasmo.* Droga A jest lepsza od B jeśli dysponuje takim samym pasmem, ale jest krótsza lub dysponuje większym pasmem i jest co najwyżej o jeden węzeł dłuższa od B.

W przypadku dostępności kilku równoważnych dróg, o wyborze decyduje mniejszy numer pierwszego różnego łącza, np. droga 1-5-6-3 jest preferowana przed 1-7-2-3.

Komentarza wymaga określenie dostępnego pasma. Jest to pasmo dostępne jednocześnie na całej długości drogi połączeniowej, ale tylko w momencie podejmowania decyzji. Po utworzeniu ścieżki następuje rezerwacja pasma, co ma wpływ na pasmo dostępne dla następnych w kolejności ścieżek.

4. WARUNKI BADAŃ

Do celów badań przygotowano sieć składającą się z 15 węzłów połączonych jak na rys. 2 łączami 155Mb/s. Topologię opracowano w taki sposób, aby każdy z węzłów posiadał identyczną liczbę wejść i wyjść. Powinno to zapobiegać sytuacji natłoku w węźle i sprzyjać możliwości wyboru dróg alternatywnych.



Rys.2. Struktura badanej sieci.

Do każdego z węzłów brzegowych 1, 2 i 3 dołączono po 3 identyczne źródła CBR i 3 identyczne źródła VBR i każde skierowano do innego węzła docelowego. Aby wyeliminować w procesie multipleksowania źródeł straty zaraz na wejściu do sieci, łącza dostępowe mają przepustowość 625Mb/s, a bufony w multiplekserach na tych łączach mają pojemność 20000 bajtów. Bufory w portach wyjściowych węzłów sieci mają pojemność 8000 bajtów dla klasy CBR i 64000 bajtów dla klasy VBR.

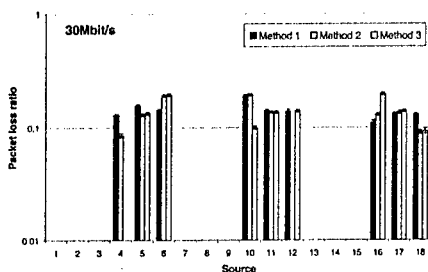
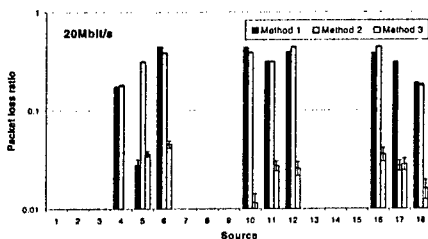
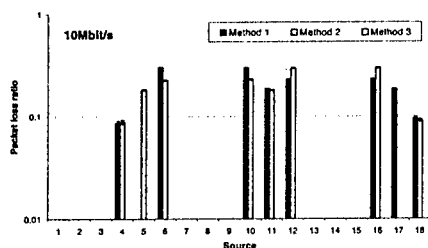
Źródła VBR generują pakiety o maksymalnej i średniej długości odpowiednio 1500 i 700 bajtów. Źródła CBR wysyłają pakiety o stałej długości 100 bajtów. Wykonano 3 serie

pomiarów, dla szybkości źródeł CBR równej 10Mb/s, 20Mb/s i 30Mbit/s. Szybkości źródeł VBR pozostawały nie zmienione i wynosiły średnio 60Mb/s i maksymalnie 99Mb/s.

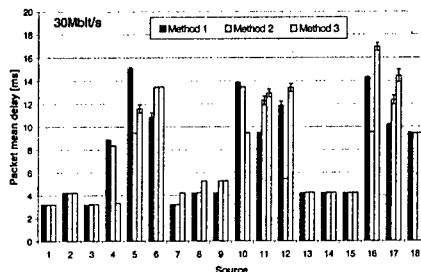
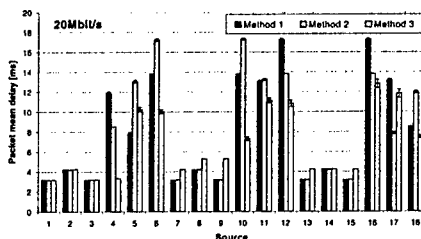
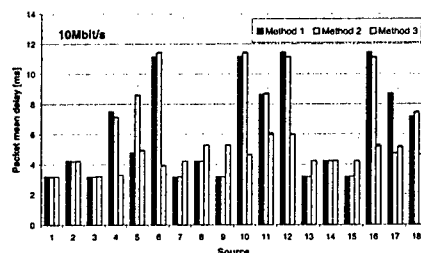
Badania symulacyjne wykonano dla 3 zdefiniowanych metod routingu, rejestrując obciążenie łączy oraz straty i średnie opóźnienie pakietów na całej długości drogi połączeniowej z podziałem na źródła generujące poszczególne strumienie. Wykonano po 20 odcinków pomiarowych symulacji po 0,2 sekundy czasu pracy rzeczywistej sieci. Przedziały ufności zostały wyliczone z rozkładu T-Studenta dla poziomu ufności 0,99.

5. WYNIKI BADAŃ

Uzyskane wyniki jednoznacznie wskazują na zależność jakości metod od obciążenia sieci. Przy mniejszym obciążeniu bardzo dobre wyniki daje metoda 3, która kosztem dłuższej drogi połączeniowej wybiera łączy mniej obciążone. Jednak przy pewnym poziomie obciążenia nie traci swoje zalety. Pozostałe dwie metody uzyskują podobne wyniki i trudno wskazać metodę zdecydowanie lepszą. Sprawdzono również, iż dalsze zwiększanie ruchu generowanego przez źródła CBR np. do 40Mb/s powoduje tylko zwiększenie strat bez zmiany charakterystyki metod.



Rys.3. Straty pakietów dla różnych szybkości źródeł CBR.

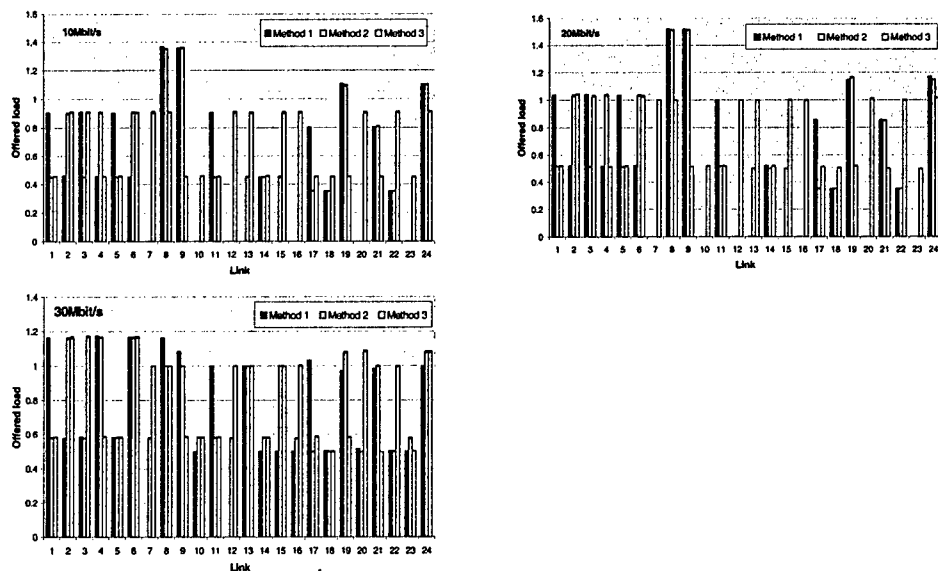


Rys.4. Średnie opóźnienie pakietów dla różnych szybkości źródeł CBR.

Wykresy opóźnień potwierdzają dobrą pracę metody 3 dla niedużego obciążenia sieci. Mimo, iż czasem wybierane są drogi dłuższe, są one mniej obciążone i dzięki temu straty i opóźnienia są mniejsze. Szczególnie dobrze jest to widoczne dla źródeł VBR, dla których bufor są dłuższe. Klasa CBR jest obsługiwana w pierwszej kolejności, dlatego w tym przypadku lepsze są metody wybierające najkrótszą drogę. Przy założeniu równej długości łączy fizycznych większa liczba węzłów pośrednich oznacza dłuższy czas propagacji, który jest znaczny dla długich łączy.

Brak strat przy małym obciążeniu dla metody 3 musi oznaczać, iż na żadnym łączy używanym przez tę metodę nie nastąpiło przeciążenie. Potwierdzają to wyniki pomiaru ruchu oferowanego. W przeciwieństwie do pozostałych metod, na żadnym z łączy ruch oferowany nie był większy od przepustowości łączy. Użycie metody 3 skutkuje uzyskaniem najbardziej równomiernego rozłożenia obciążenia. Zasada ta nie stosuje się już jednak w przypadku dużego obciążenia sieci, gdyż kluczem do jego osiągnięcia jest wybieranie dłuższych łączy, co skutkuje sumarycznie większą ilością ruchu w sieci.

Warto zwrócić uwagę, że metody 1 i 2 nie wykorzystują w pełni możliwości sieci pozostawiając pewne łączy nieużywane. Metoda 1 nawet przy największym badanym obciążeniu nie korzysta w ogóle z łączy 7 i 12. Prowadzi to do większego obciążenia i strat na pozostałych łączyach.



Rys.5. Średni ruch oferowany na łączy dla różnych szybkości źródeł CBR.

Podsumowując, dla założonej topologii metoda 3 okazała się najlepsza, co zostało potwierdzone dobrymi wynikami strat i opóźnień pakietów. Należy zaznaczyć, iż takie wyniki nie mogą być uogólniane na wszystkie możliwe topologie. Jeśli sieć nie sprzyja wybieraniu dróg alternatywnych, wówczas metoda 3 nie musi okazać się lepsza od pozostałych [6].

6. PODSUMOWANIE

Wykonany program wykazał swoją przydatność w symulowaniu sieci MPLS. Jego elastyczność pozwala w stosunkowo łatwy sposób zdefiniować sieć i postać raportów. Uzyskano interesujące wyniki, wskazujące na silną zależność pomiędzy warunkami ruchowymi w sieci a rezultatami uzyskanymi za pomocą badanych metod kierowania ruchem. Dla wybranej topologii najlepszą okazała się metoda 3, która przedkłada wolne pasmo nad długością drogi. Czynnikiem sprzyjającym jest tu zrównoważona topologia sieci.

Dalsze kierunki badań i implementacji będą obejmowały bardziej wyrafinowane metody routingu, szczególnie *explicit routing* i metoda dynamiczna.

BIBLIOGRAFIA

- [1] Kaczmarek S., *MPLS i inżynieria ruchu*, W: mat. konf. INTERNET – Wrocław 2002, Grudzień 2002, 9-18.
- [2] Awduche D. i in., *Requirements for Traffic Engineering over MPLS*, RFC 2702, Sept. 1999.
- [3] Callon R., Rosen E., Viswanathan A., *Multiprotocol Label Switching Architecture*, RFC 3031, Jan. 2001.
- [4] Blake S. i in., *An Architecture for Differentiated Service*, RFC2475. Dec. 1998.
- [5] Ahmadi H., Guérin R., Naghshineh M., *Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks*, IEEE JSAC, Sept. 1991, no. 7, vol. 9, 968-981.
- [6] Kaczmarek S., Nowak K., *Symulacja pracy sieci MPLS*, W: mat. konf. INTERNET – Wrocław 2003, Grudzień 2003, 102-112.

SIMULATION MODEL AND STUDY OF MPLS NETWORK

Summary

Multiprotocol Label Switching (MPLS) is a promising technique providing support for quality of service in core networks based on IP protocol. To be able to take an advantage of it, one needs to develop efficient algorithms for QoS guarantees and one of the key issues here is routing. A necessary tool for such purposes is simulation program. This paper presents features of such program and describes the implementation. Sample 15-node MPLS network was presented. Based on simulation results, three simple routing methods were compared and some conclusions were drawn. Finally directions for further studies in that area were presented.

Adrian Kosowski

**Katedra Algorytmów i Modelowania Systemów
Wydział Elektroniki, Telekomunikacji i Informatyki
Politechnika Gdańska**

WYBRANE WŁASNOŚCI PROBLEMU ROUTINGU ORAZ KOLOROWANIA ŚCIEŻEK W GRAFIE

Streszczenie

Referat dotyczy zagadnienia ścieżkowego kolorowania grafu, które stanowi naturalny model dla problemu routingu i przydziału częstotliwości w czysto optycznej sieci światłowodowej. Opisano podstawowe zasady i właściwości ścieżkowego kolorowania grafów. Zaprezentowano wybrane twierdzenia, oparte w dużej mierze na wynikach badań własnych. Omówiono złożoność obliczeniową problemu routingu chromatycznego i kolorowania ścieżek zarówno w przypadku ogólnym, jak i dla wybranych klas grafów. Scharakteryzowano złożoność obliczeniową badanych problemów w przypadkach szczególnych, gdy ograniczony jest albo rząd grafu, albo ograniczona jest moc zbioru ścieżek.

1. WPROWADZENIE

Istnieją liczne problemy optymalizacyjne w zakresie projektowania i zautomatyzowanego zarządzania przepływem danych w sieciach komputerowych, które można sformułować korzystając z terminologii routingu i kolorowania ścieżek. Graf stanowi interpretację topologii sieci komputerowej, przy czym wierzchołki grafu oznaczają węzły sieci, a krawędzie — połączenia pomiędzy węzłami. Jeżeli połączenia między węzłami sieci są jednokierunkowe, sieć należy modelować jako digraf z łukami o zwrocie odpowiadającym kierunkowi łącza.

Modele kolorowania ścieżek w grafie odgrywają największą rolę w sieciach opartych na rozwiązaniach optycznych, najczęściej nie wymagających elektronicznej konwersji sygnału optycznego pomiędzy węzłem początkowym (nadawcą) a końcowym (odbiorcą). Obecnie, łącza optyczne (światłowody) są w stanie realizować transmisję jednokierunkową danych, z możliwością zwielokrotnienia pasma poprzez równoległą transmisję używając fal optycznych o różnych częstotliwościach (od 4 do 50 wartości częstotliwości, w zależności od typu światłowodu [1]). Z tego względu topologię sieci optycznej modeluje się najczęściej w postaci digrafu. Urządzenia pozwalające na konwersję częstotliwości sygnału optycznego w węzle pośrednim transmisji, pomimo przeszło 10 lat rozwoju technologii, zazwyczaj stosują mechanizm rozkodowywania informacji optycznej do postaci elektro-

nicznej i ponownego jej kodowania jako fali o innej częstotliwości. Takie działanie prowadzi nie tylko do opóźnień transmisji, ale może zmniejszyć prędkość komunikacji z powodu ograniczonej szybkości działania konwerterów optyczno-elektronicznych i elektroniczno-optycznych, stając się „wąskim gardłem” dla łącz optycznych o potencjalnie dużej przepustowości. Dążenie do eliminacji lub choćby minimalizacji liczby konwerterów częstotliwości prowadzi do poważnych problemów optymalizacyjnych nie tylko na etapie konstrukcji sieci, ale także przy zarządzaniu trasami i parametrami transmisji danych w sieci o znanej topologii.

Praca poświęcona jest problemom routingu i przydziału częstotliwości w tzw. sieciach *czysto optycznych*, w których węzły sieci nie są wyposażone w konwertery częstotliwości fali, a jedynie w programowalne urządzenia przełączające, pozwalające na sterowanie dalszym biegiem fali świetlnej (ang. *switching*) w zależności od jej częstotliwości. Bardziej szczegółowy opis modelu sieci czysto optycznych oraz modeli pokrewnych znaleźć można w pracach [2], [3], [4].

Nawiązanie łączności pomiędzy dwoma węzłami u, v wymaga ustalenia częstotliwości fali świetlnej używanej przy transmisji danych oraz trasy transmisji danych, tj. pewnej ścieżki łączącej wierzchołki u i v w digrafie reprezentującym topologię sieci. Ustalenie trasy transmisji odbywa się w sposób fizyczny poprzez elektroniczne zaprogramowanie przełączników w poszczególnych węzłach na ścieżce łączącej u i v . Trasa oraz częstotliwość fali muszą być tak dobrane, aby transmisje po kolidujących ścieżkach nie wykorzystywały tych samych częstotliwości.

Jak już wcześniej wspomniano, ze względu na jednokierunkowy charakter łącz światłowodowych, połączenia pomiędzy węzłami sieci interpretuje się jako łuki digrafu opisującego topologię sieci. Model światłowodów dwukierunkowych, w którym topologia sieci jest opisana za pomocą grafu nieskierowanego, zostanie również opisany ze względu na silny związek pomiędzy wynikami uzyskiwanymi przy analizie grafów oraz digrafów.

2. PODSTAWOWE POJĘCIA

Rozważania o kolorowaniu ścieżek należy rozpocząć od formalnego zdefiniowania pojęcia ścieżki.

Definicja 1. Niech będzie dany graf prosty G o zbiorze wierzchołków V i zbiorze krawędzi E , $G = (V, E)$. Ciąg krawędzi (e_1, e_2, \dots, e_k) grafu G nazywamy *ścieżką* długości k , jeżeli dla pewnych wierzchołków v_1, v_2, \dots, v_{k+1} grafu G zachodzi: $e_1 = \{v_1, v_2\}$, $e_2 = \{v_2, v_3\}$, ..., $e_k = \{v_k, v_{k+1}\}$. Wierzchołki v_1 i v_{k+1} nazywane są wówczas *końcami ścieżki*. Symbol $P\{u, v\}$ używany jest do oznaczenia dowolnej ścieżki o końcach u i v . Analogicznie, w przypadku digrafu prostego $D = (V, E)$, ciąg łuków (e_1, e_2, \dots, e_k) nazywany jest *ścieżką (skierowaną)*, jeśli dla pewnych wierzchołków v_1, v_2, \dots, v_{k+1} digrafu D zachodzi: $e_1 = (v_1, v_2)$, $e_2 = (v_2, v_3)$, ..., $e_k = (v_k, v_{k+1})$. Wierzchołek v_1 jest wówczas *początkiem* ścieżki skierowanej, a v_{k+1} – jej *końcem*. Ścieżkę o początku u i końcu v oznaczamy $P(u, v)$.

Przy omawianiu problemów związanych z routingiem oraz przydziałem częstotliwości wygodnie jest posługiwać się sformalizowanym pojęciem pary ścieżek kolidujących (tzn. ścieżek transmisji, którym nie może zostać przydzielona ta sama częstotliwość).

Definicja 2. Ścieżki P_1 i P_2 w danym grafie (lub digrafie) nazywamy *kolidującymi*, gdy nie są krawędziowo rozłączne, tzn. $\exists_{e \in E} : (e \in P_1 \wedge e \in P_2)$.

Definicja 3. Dla danego zbioru ścieżek R w grafie G (w digrafie D), zbiór par ścieżek kolidujących oznaczamy $C(R)$. Dla zbioru ścieżek R w grafie G można jednoznacznie określić *graf konfliktów* $Q(R)$ jako graf, którego zbiorem wierzchołków jest R , a krawędź między dwoma wierzchołkami istnieje, gdy odpowiadające im ścieżki grafu G są kolidujące: $Q(R) = (R, C(R))$.

Innym istotnym parametrem jest obciążenie grafu przez zbiór ścieżek, które można interpretować jako miarę stopnia wykorzystania przepustowości łącz w sieci komputerowej.

Definicja 4. Obciążenie $L_e(R)$ krawędzi e grafu G (łuku e digrafu D) przez zbiór ścieżek R w tym grafie (digrafie) definiujemy jako liczbę ścieżek ze zbioru R , które zawierają krawędź (łuk) e . Maksymalne obciążenie krawędzi w grafie G (digrafie D) przez zbiór ścieżek R nazywane jest *obciążeniem grafu (digrafu)* i oznaczane jest symbolem $L_G(R)$ ($L_D(R)$).

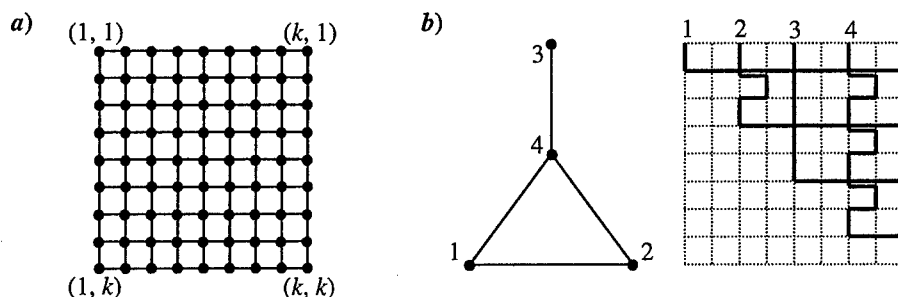
3. WŁASNOŚCI KOLOROWANIA ŚCIEŻEK

Definicja 5. Pokolorowaniem zbioru ścieżek R za pomocą k kolorów (lub k -pokolorowaniem zbioru R) nazywamy takie przyporządkowanie dokładnie jednej z liczb $1, 2, \dots, k$ każdej ścieżce ze zbioru R , że ścieżki kolidujące otrzymują różne kolory.

Nietrudno zauważyć, że kolorowanie zbioru ścieżek R jest równoważne klasycznemu kolorowaniu wierzchołkowemu grafu konfliktów $Q(R)$ odpowiadającego zbiorowi R . Faktycznie, niech $R = \{P_1, P_2, \dots, P_l\}$, wówczas funkcja $c: R \rightarrow \{1, 2, \dots, k\}$ jest k -pokolorowaniem zbioru R , jeśli $\forall_{\{P_i, P_j\} \in C(R)}: c(P_i) \neq c(P_j)$.

Definicja 6. Najmniejszą liczbę naturalną k , dla której istnieje k -pokolorowanie zbioru ścieżek R w grafie G (digrafie D) nazywamy *liczbą chromatyczną* zbioru R i oznaczamy symbolem $\chi_G(R)$ ($\chi_D(R)$) lub, gdy nie prowadzi to do niejasności, $\chi(R)$. Liczba chromatyczna zbioru R jest równa klasycznej liczbie chromatycznej grafu konfliktów $Q(R)$.

Twierdzenie 1 [5]. Problem znajdowania wartości liczby chromatycznej $\chi_G(R)$ jest trudny obliczeniowo, nawet dla grafów G o topologii kratownicy (rysunek 1a).



Rys.1. a) Graf o topologii kratownicy kwadratowej z numeracją współrzędnych wierzchołków, b) Przykład grafu czterowierzchołkowego i skonstruowanego zbioru ścieżek w kratownicy kwadratowej 9×9 o tej samej liczbie chromatycznej co zadany graf (numeracja ścieżek odpowiada numeracji wierzchołków w grafie, ścieżki zawierające wspólną krawędź są dla większej czytelności rozsunięte w pionie)

Dowód twierdzenia 1 ma charakter konstrukcyjny i polega na sprowadzeniu (redukcji) problemu znajdowania klasycznej liczby chromatycznej grafu do problemu znajdowania liczby chromatycznej pewnego zbioru ścieżek w kratownicy. Dla danego grafu $G = (V, E)$ rzędu n , konstruujemy kratownicę kwadratową M o $(2n+1)^2$ węzłach oraz zbiór R złożony z n ścieżek w kratownicy M , taki że $\chi_M(R) = \chi(G)$. Niech zbiór wierzchołków będzie dany jako $V = \{v_1, v_2, \dots, v_n\}$, zbiór ścieżek $R = \{P_1, P_2, \dots, P_n\}$. Ścieżkę P_i można skonstruować zgodnie z następującym algorytmem¹:

1. rozpocznij ścieżkę P_i w wierzchołku kratownicy o współrzędnych $(2i-1, 1)$ i dołącz do ścieżki krawędź odpowiadającą przemieszczeniu końca ścieżki o wektor $(0, 1)$;
2. dla $j = 1, 2, \dots, i-1$ wykonaj:
 - a. jeśli $\{v_i, v_j\} \in E$, to do ścieżki P_i dołącz cztery krawędzie, odpowiadające kolejnym przemieszczeniom końca ścieżki o wektory $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, 1)$;
 - b. w przeciwnym wypadku, do ścieżki P_i dołącz dwie krawędzie, odpowiadające kolejnym przemieszczeniom końca ścieżki o wektory $(0, 1)$, $(0, 1)$;
3. do końca ścieżki P_i dołącz $2n-2i+2$ krawędzi, przemieszczając koniec ścieżki za każdym razem o wektor $(1, 0)$.

Łatwo pokazać, że w tak skonstruowanym zbiorze ścieżek R ścieżki P_i, P_j kolidują wtedy i tylko wtedy, gdy wierzchołki v_i, v_j sąsiadują w grafie G . Zatem $G = Q(R)$, co kończy dowód. Przykład grafu G i odpowiadającego mu zbioru ścieżek R w kratownicy M przedstawiono na rysunku 1b.

W ogólności nie jest znany efektywny (tj. deterministyczny wielomianowy) algorytm k -względnie przybliżony pozwalający wyznaczyć $\chi_G(R)$ dla zadanego zbioru ścieżek R w zadanym grafie lub digrafie G . Algorytmy takie istnieją, jeśli graf G jest drzewem albo cyklem nieskierowanym. Jeśli graf G jest *drogą* (drzewem o dwóch liściach), to istnieje efektywny algorytm wyznaczający $\chi_G(R)$ dla każdego zbioru ścieżek R .

Twierdzenie 2. Niech $G = (V, E)$ będzie dowolnym grafem lub digrafem, a R zbiorem ścieżek w G . Zachodzą następujące oszacowania liczby chromatycznej $\chi_G(R)$:

$$\chi_G(R) = \chi(Q(R)) \geq \omega(Q(R)) \geq L_G(R) \quad (3.1)$$

$$\chi_G(R) \leq 1 + \Delta(Q(R)) \leq 1 + (L_G(R) - 1) \cdot \max_{P \in R} |P| \quad (3.2)$$

$$\chi_G(R) \leq L_G(R) \cdot \sqrt{|E|} \quad (3.3)$$

gdzie $Q(R)$ oznacza graf konfliktów zbioru ścieżek R , $|P|$ oznacza długość ścieżki P , $|E|$ oznacza moc zbioru E , $\omega(Q(R))$ to maksymalny rząd podgrafu pełnego grafu $Q(R)$ (liczba klikowa grafu $Q(R)$), a $\Delta(Q(R))$ to stopień tego grafu (maksymalna liczba sąsiadów wierzchołka grafu $Q(R)$).

Dowód oszacowania (3.1) jest natychmiastowy, ponieważ ścieżki biegnące wzdłuż najbardziej obciążonej krawędzi grafu G indukują podgraf pełny w grafie konfliktów $Q(R)$,

¹ Prezentowana tu konstrukcja jest istotnie prostsza niż ta zastosowana w oryginalnym dowodzie [5] i wykorzystuje kratownicę M o mniejszym rzędzie. Pewne własności tej konstrukcji są istotne przy dalszych rozważaniach.

natomiast zależność pomiędzy liczbą chromatyczną a liczbą klikową grafu $Q(R)$ jest prawdziwa dla każdego grafu. Oszacowanie (3.2) wynika z faktu, że żadna ścieżka nie może kolidować z więcej niż $(L_G(R)-1) \cdot \max_{P \in R} |P|$ innymi ścieżkami [6]. Aby uzasadnić oszacowanie (3.3), rozpatrzmy osobno ścieżki o długości nie mniejszej niż $|E|^{1/2}$ oraz ścieżki krótsze. Ścieżek o długości nie mniejszej niż $|E|^{1/2}$ jest nie więcej niż $L_G(R) \cdot |E|^{1/2}$, gdyż obciążenie żadnej z krawędzi nie może przekraczać $L_G(R)$. W najbardziej pesymistycznym przypadku ścieżki te tworzą w grafie konfliktów $Q(R)$ klikę rzędu $L_G(R) \cdot |E|^{1/2}$. Każda z pozostałych (krótszych) ścieżek koliduje z co najwyżej $(L_G(R)-1) \cdot |E|^{1/2}$ innymi ścieżkami, a więc jej stopień w grafie konfliktów $Q(R)$ nie przekracza $(L_G(R)-1) \cdot |E|^{1/2}$. Kolorując graf konfliktów $Q(R)$ można przyporządkować kolory $1, 2, \dots, L_G(R) \cdot |E|^{1/2}$ wierzchołkom wspomnianej kliki. Po przydzieleniu każdemu z pozostałych wierzchołków (o stopniu ograniczonym przez $(L_G(R)-1) \cdot |E|^{1/2}$) najmniejszego dopuszczalnego koloru, uzyskuje się pokolorowanie grafu $Q(R)$ przy pomocy $L_G(R) \cdot |E|^{1/2}$ kolorów, zatem $\chi(Q(R)) \leq L_G(R) \cdot |E|^{1/2}$, co należało pokazać. Oszacowanie (3.3) jest dwukrotnie silniejsze od oszacowania zaproponowanego w pracach [6] i [7].

4. PROBLEM ROUTINGU

Definicja 7. *Zgłoszeniem* w grafie nieskierowanym nazywamy dowolną wyróżnioną parę (u, v) wierzchołków tego grafu. Zgłoszenie w digrafie definiujemy analogicznie, przy czym istotne jest uporządkowanie pary (u, v) .

Definicja 8. *Instancja* I w grafie G (digrafie D) zdefiniowana jest jako dowolny zbiór zgłoszeń w tym grafie (digrafie), z dopuszczeniem możliwości powtarzania się elementów.

Definicja 9. *Routingiem* $R(I)$ instancji I w grafie nieskierowanym G nazywamy zbiór ścieżek w G , taki że I jest zbiorem końców ścieżek z $R(I)$. Analogicznie, dla instancji I w digrafie D routing $R(I)$ definiujemy jako zbiór ścieżek w D , taki że I jest zbiorem uporządkowanych par wierzchołków początkowych i końcowych dla ścieżek z $R(I)$. Również w przypadku routingu dopuszcza się możliwość powtarzania się elementów zbioru (tj. ta sama ścieżka może występować w routingu kilkakrotnie).

Definicja 10. *Liczbę chromatyczną* $\chi(I)$ instancji I definiuje się jako najmniejszą możliwą liczbę chromatyczną routingu instancji I : $\chi(I) = \min_{R(I)} \chi(R(I))$.

Przy poszukiwaniu optymalnego routingu dla instancji I , zazwyczaj przyjmuje się za kryterium optymalizacji dyskretnej minimalizację liczby kolorów koniecznych do pokolorowania routingu $R(I)$. W języku teorii grafów formułuje się zazwyczaj trzy podstawowe problemy optymalizacyjne (podane poniżej w ujęciu dla grafów nieskierowanych):

CHROMATICROUTING. Dla zadanej instancji I w grafie G znaleźć routing $R(I)$, dla którego liczba chromatyczna $\chi_G(R(I))$ jest równa $\chi_G(I)$.

PATHCOLORING. Dlaadanego zbioru ścieżek R w grafie G znaleźć pokolorowanie o rozpiętości $\chi_G(R)$ kolorów.

CHROMATICROUTINGANDPATHCOLORING. Dla zadanej instancji I w grafie G znaleźć routing $R(I)$, dla którego liczba chromatyczna $\chi_G(R(I))$ jest równa $\chi_G(I)$ oraz wyznaczyć pokolorowanie tego routingu o rozpiętości $\chi_G(I)$.

5. ANALIZA ZŁOŻONOŚCI OBLICZENIOWEJ PROBLEMÓW ROUTINGU I KOLOROWANIA ŚCIEŻEK

Wszystkie wymienione w poprzednim punkcie problemy optymalizacyjne są w przypadku ogólnym *NP*-trudne, zarówno w ujęciu dla grafów nieskierowanych, jak i digrafów ([6], [7], [8]). Problemy *PATHCOLORING* i *CHROMATICROUTING* należą w ogólności do klasy *APX*. Problem *PATHCOLORING* pozostaje *NP*-trudny nawet dla grafów nieskierowanych o topologii gwiazdy (w tym szczególnym przypadku problem jest równoważny kolorowaniu krawędziowemu multigrafu, zatem istnieje dla niego algorytm aproksymacyjny ze współczynnikiem dobroci 1.1, [9]).

Zwraca uwagę pozornie zaskakujący fakt, że problem powstający z połączenia problemu routingu i kolorowania ścieżek (*CHROMATICROUTINGANDPATHCOLORING*) może okazać się w pewnych przypadkach łatwiejszy w aproksymacji, niż sam problem kolorowania ścieżek (*PATHCOLORING*). Za przykład może tu posłużyć klasa grafów o topologii kratownicy prostokątnej. W [10] pokazano, że dla problemu routingu i kolorowania ścieżek dla instancji złożonej z n zgłoszeń w kratownicy o bokach długości $\theta(n)$ istnieje algorytm aproksymacyjny prowadzący do rozwiązania wykorzystującego co najwyżej $O(\text{poly}(\log \log n))$ razy więcej kolorów niż rozwiązanie optymalne (poly oznacza tu pewną ustaloną funkcję wielomianową). Z drugiej strony, konstrukcja z dowodu twierdzenia 1 pokazuje, że problemu kolorowania zbioru n ścieżek w takiej samej kratownicy prostokątnej jest równoważny kolorowaniu klasycznemu grafów, zatem żaden algorytm przybliżony o podobnej funkcji dobroci nie jest znany.

Na uwagę zasługują również charakterystyczne przypadki, w których problemy *CHROMATICROUTING* i *PATHCOLORING* mogą się różnić nawet klasą złożoności obliczeniowej.

Twierdzenie 3 [6]. *Dla grafu nieskierowanego G będącego drzewem problem *CHROMATICROUTING* jest rozwiązywalny wielomianowo, podczas gdy problem *PATHCOLORING* jest *NP*-trudny i istnieje dla niego algorytm aproksymacyjny o współczynniku dobroci 1.1.*

Twierdzenie 4. *Niech $k > 1$ będzie ustaloną liczbą naturalną. Problem *PATHCOLORING* dla zbioru k ścieżek w grafie G jest rozwiązywalny wielomianowo. Problem *CHROMATICROUTING* dla instancji złożonej z k zgłoszeń jest *NP*-trudny nawet dla $k=2$.*

Wielomianowy algorytm rozwiązujący problem *PATHCOLORING* dla zbioru k ścieżek wyznacza w czasie $O(n)$ k -wierzchołkowy graf konfliktów dla tego zbioru ścieżek, po czym koloruje go w czasie $O(1)$. Problem *CHROMATICROUTING* jest *NP*-trudny, ponieważ rozstrzygnięcie, czy dla danej instancji $\{(u_1, v_1), (u_2, v_2)\}$ złożonej z wierzchołków grafu G istnieje para niekolidujących ścieżek $P\{u_1, v_1\}, P\{u_2, v_2\}$ w grafie G , jest *NP*-zupełne (dowód przebiega analogicznie jak dowód *NP*-zupełności problemu *DISJOINTCONNECTEDPATHS* ND40, opisany w [11]).

Twierdzenie 4 jest również prawdziwe dla routingu i kolorowania ścieżek w digrafach.

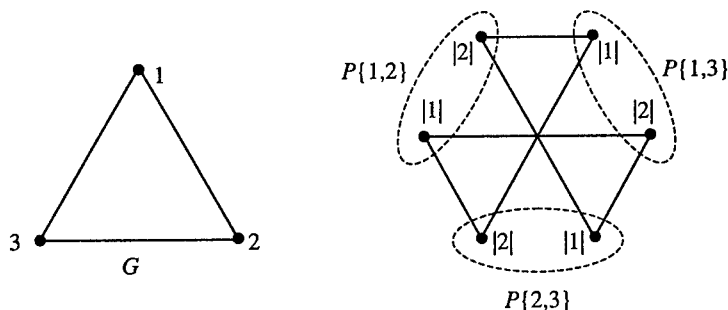
Twierdzenie 5. *Niech $k > 1$ będzie ustaloną liczbą naturalną. Problemy *PATHCOLORING*, *CHROMATICROUTING*, *CHROMATICROUTINGANDPATHCOLORING* można rozwiązać w czasie wielomianowym (względem mocy instancji czy zbioru ścieżek) dla grafu G rzędu k .*

W dowodzie twierdzenia 5 przydatny jest następujący lemat.

Lemat. Niech będzie dany graf $G = (V, E)$ ustalonego rzędu k oraz funkcja $x: V \rightarrow \mathbb{N}$, taka że $\sum_{v \in V} x(v) = n$. Wówczas optymalne multipokolorowanie grafu G (tj. przyporządkowanie każdemu wierzchołkowi $v \in V$ zbioru mocy $x(v)$ liczb naturalnych z przedziału $\{1, 2, \dots, c\}$, tak aby przekrój zbiorów przyporządkowanych sąsiadującym wierzchołkom był pusty, a c przyjmowało możliwie małą wartość naturalną, [12]) można wyznaczyć w czasie $O(\text{poly}(n))$.

Prawdziwość lematu wykażemy poprzez wskazanie wielomianowego algorytmu prowadzącego do odpowiedniego multipokolorowania. Kolorując zachłannie otrzymuje się pokolorowanie o rozpiętości nie przekraczającej n , zatem $c \leq n$. Pozostaje do pokazania, że można w czasie wielomianowym dla ustalonego $d \leq n$ wskazać d -multipokolorowanie grafu G , lub stwierdzić, że takie pokolorowanie nie istnieje. Każde d -multipokolorowanie można interpretować jako pewną funkcję przypisującą każdemu elementowi i zbioru kolorów $\{1, 2, \dots, d\}$ dokładnie jeden podzbiór V_i zbioru wierzchołków V , którego elementami są wszystkie wierzchołki, w których występuje kolor i . Równoważnie, można zaindeksować podzbiory zbioru V liczbami naturalnymi $\{1, 2, \dots, 2^k\}$ i traktować d -multipokolorowanie jako pewną funkcję f o argumentach ze zbioru $\{1, 2, \dots, d\}$ i wartościach w zbiorze $\{1, 2, \dots, 2^k\}$. Ponieważ przypisanie kolorom etykiet liczbowych jest arbitralne, można bez utraty ogólności ograniczyć się wyłącznie do tych pokolorowań, których odwzorowania f są niemalejące. Wystarczy więc, aby algorytm sprawdzał, czy w zbiorze wszystkich odwzorowań niemalejących ze zbioru $\{1, 2, \dots, d\}$ w zbiór $\{1, 2, \dots, 2^k\}$ istnieje legalne d -multipokolorowanie grafu G . Ponieważ $d \leq n$, złożoność rozpatrywanego algorytmu jest wielomianowa względem n . Prawdziwość lematu została udowodniona.

Pokażemy teraz wielomianową złożoność obliczeniową problemu PATHCOLORING dla zbioru ścieżek R w grafie G ustalonego rzędu k . Niech H oznacza graf konfliktów dla zbioru wszystkich możliwych różnych ścieżek w grafie G (rysunek 2). Rząd grafu H nie przekracza pewnej stałej wartości (mniejszej od k^k). Każdemu wierzchołkowi P grafu H przyporządkowuje się wartość $x(P)$ odpowiadającą liczbie wystąpień ścieżki P w zbiorze ścieżek R . Multipokolorowanie grafu H z funkcją x prowadzi do jednoznacznej odpowiedzi na problem kolorowania zbioru ścieżek R w czasie wielomianowym.



Rys.2. Przykład trójkątnego grafu $G=C_3$ i odpowiadającego mu grafu konfliktów dla zbioru wszystkich różnych ścieżek (symbole podawane przy wierzchołkach grafu konfliktów oznaczają długości ścieżek w G)

Pozostała do udowodnienia wielomianowa złożoność obliczeniowa problemów CHROMATICROUTING i CHROMATICROUTINGANDPATHCOLORING dla grafu G ustalonego rzędu k . Liczba istotnie różnych zgłoszeń w instancji I w grafie G nie może przekroczyć stałej wartości $k \cdot (k-1)/2$. Liczba różnych ścieżek łączących dowolne dwa wierzchołki w grafie G jest ograniczona od góry przez stałą wartość k^k . Zatem liczba różnych routingu dla instancji I złożonej z n zgłoszeń jest wielomianowa względem n , $O(\text{poly}(n))$. W połączeniu z uzyskanymi w poprzednim paragrafie wynikami dla problemu PATHCOLORING, można bez trudu wywnioskować, że w rozważanym przypadku omawiane problemy routingu i kolorowania są wielomianowo rozwiązywalne.

BIBLIOGRAFIA

- [1] Beauquier B., Bermond J.C., Gargano L., Hell P., Perennes S., Vaccaro U.: *Graph problems arising from wavelength routing in all-optical networks*. W: Proc. 2nd Workshop on Optics and Computer Science, 1997, Paryż, Francja.
- [2] Raghavan R., Upfal E.: *Efficient routing in all-optical networks*. W: Proc. 26th ACM Symposium on the Theory of Computing, 1994, s. 134–143.
- [3] Gargano L., Vaccaro U.: *Routing in all-optical networks: algorithmic and graph-theoretic problems*. W: Numbers, Information and Complexity, Kluwer, 2000, Boston, USA, s. 555–578.
- [4] Green P.E.: *Fiber-optic communication networks*. Prentice-Hall, 1992.
- [5] Nomikos C.: *Path coloring in graphs*. Praca doktorska, Dept. of Electrical and Computer Engineering, NTUA, 1997, Ateny, Grecja.
- [6] Białogrodzki J.: *Kolorowanie ścieżek w grafach*. W: Modele i metody kolorowania grafów, Kubale M. et al, WNT, 2002, Warszawa, s. 190–208.
- [7] Aggrawal A., Bar-Noy A., Coppersmith D., Ramsawami R., Schieber B., Sudan M.: *Efficient routing and scheduling algorithms for optical networks*. W: Proc. 5th Annual ACM-SIAM Symposium on Discrete Algorithms, 1994, s. 412–423.
- [8] Erlebach T., Jansen K.: *The complexity of path coloring and call scheduling*. W: Theoretical Computer Science, Elsevier vol. 255, 2001, s. 33–50.
- [9] Erlebach T., Jansen K., Kaklamani C., Persiano P.: *An optimal greedy algorithm for wavelength allocation in directed tree networks*. W: Proc. DIMACS Workshop on Network Design, 1997.
- [10] Rabani Y.: *Path Coloring on the Mesh*. W: Proc. IEEE Symposium on Foundations of Computer Science, 1996, s. 400–409.
- [11] Srinivasan, A.: *Improved approximations for edge-disjoint paths, unsplittable flow, and related routing problems*. Proc. 38th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 1997, s. 416–425.
- [12] Halldorsson M.M., Kortsarz G.: *Multicoloring Planar Graphs and Partial k -Trees*. W: Proc. Random-Approx Conference, 1997, s. 73–84.

SOME ASPECTS OF ROUTING AND PATH COLORING IN GRAPHS

Summary

The paper deals with the issue of coloring paths in a graph, which may be regarded as the most natural model for the request routing and frequency assignment problems in an all-optical fiber-optic network. Basic properties and principles of path coloring are described, including some chosen theorems, based largely on results of the author's research. The computational complexity of the chromatic routing problem and the path coloring problem is discussed both in the general case and for certain special classes of graphs. A complete analysis of the computational complexity of the problems is performed for the case of graphs of bounded order and the case of sets of paths of bounded order.

Leszek Mazur, Artur Stencel, Daniel R. Świeciak, Ryszard Weisbrodt

DGT Sp. z o.o.,

Katedra Systemów i Sieci Telekomunikacyjnych, Politechnika Gdańska

SYSTEM ZARZĄDZANIA RUCHEM W SIECI Z WĘZŁAMI DGT

Streszczenie

W referacie przedstawiono opracowany we współpracy z firmą DGT system informatyczny do zarządzania ruchem w rzeczywistej sieci telekomunikacyjnej z węzłami DGT. Referat składa się z dwóch części. W części związanej z zagadnieniami telekomunikacyjnymi przedstawiono wyniki analizy metod zarządzania ruchem w sieci pod kątem możliwości zaimplementowania ich w sieci z węzłami DGT. Dla wybranych metod zarządzania określono wymagane parametry wyznaczone w procesie monitorowania węzłów sieci. W części informatycznej przedstawiono wymagania i specyfikację funkcjonalną na system oraz sposób ich realizacji. Zgodnie z przyjętą metodologią omówiono sposób powstawania oprogramowania od koncepcji do realizacji w pełni funkcjonalnego produktu. Omówiono wyniki przeprowadzonych testów.

1. WSTĘP

Zarządzanie wydajnością sieci telekomunikacyjnej jest ważnym obszarem funkcjonalnym jej zarządzania. Udostępnia on środki do optymalnego wykorzystania zasobów sieci a celem ich stosowania jest obniżenie kosztów utrzymania sieci, co w konsekwencji sprzyja zwiększeniu zysku operatora. Jednym z podobszarów zarządzania wydajnością jest zarządzanie ruchem. We współczesnych sieciach telekomunikacyjnych realizacja funkcji zarządzania ruchem jest wspomagana przez systemy informatyczne.

Celem referatu jest przedstawienie wyniku pracy [1], którym jest system informatyczny pozwalający za pomocą wybranych metod zarządzania ruchem nadzorować i sterować przepływem strumieni ruchu w sieci telekomunikacyjnej, a w szczególności w sieci z węzłami systemu DGT. System został opracowany przy współpracy z firmą DGT i w sposób istotny uzupełnia funkcjonalność central DGT o zaimplementowane w nim metody zarządzania. Z uwagi na fakt, że system zarządzania ruchem ma współpracować z węzłami DGT i firmowym systemem zarządzania SZ – DGT, opartych na systemie operacyjnym firmy Microsoft®, przyjęto założenie, że tworzony system zarządzania będzie oparty na platformie Microsoft® Windows™. Mając jednak na uwadze zastosowanie systemu w innych środowiskach oraz przyjętą architekturę obiektową oprogramowanie systemu zarządzania wykonano w języku JAVA™. Język ten wspiera obiektowe tworzenie oprogramowania, a jego niezależność od używanego systemu operacyjnego pozwala w łatwy sposób dokonać kompilacji kodu w dowolnym środowisku (m.in. Unix/Linux).

2. LOKALIZACJA SYSTEMU ZARZĄDZANIA RUCHEM

System zarządzania ruchem w sieci lokalizowany jest w warstwie zarządzania siecią (*Network Management Layer*) a podejmowane przez ten system działania dotyczą warstwy zarządzania elementami sieci (*Network Element Management Layer*).

W warstwie zarządzania siecią podejmuje się decyzje dotyczące funkcjonowania sieci jako całości. Oznacza to m.in. wybór i przydział zasobów sieciowych, nadzorowanie stanów ruchowych i sprawności łączy, administrowanie połączeniami sieciowymi. Celem tych działań jest poprawa wydajności obsługi ruchu przenoszonego przez sieć. Poprawne funkcjonowanie tej warstwy zarządzania związane jest z prowadzeniem i utrzymaniem odpowiednich baz danych i przetwarzaniem zawartych w nich informacji.

W warstwie zarządzania elementami sieci podejmuje się decyzje dotyczące funkcjonowania poszczególnych elementów sieci. Dotyczą one m.in. konfigurowania i przydzielania zasobów sieciowych, alarmowanie o uszkodzeniach, kontrolowanie stanów elementów. Podobnie jak w warstwie zarządzania siecią tak i w tym przypadku dla poprawnego funkcjonowania tej warstwy niezbędne są bazy danych o elementach sieci, o ich konfiguracji, o uszkodzeniach itp.

Dla obu wyżej wymienionych warstw zarządzania, z punktu widzenia określenia wymagań dla opracowanego systemu zarządzania ruchem, istotne były funkcje należące do obszaru zarządzania konfiguracją (*Configuration Management*). Do głównych zadań tego obszaru należą: instalowanie i uruchamianie zasobów sieciowych, konfigurowanie zasobów, administrowanie zasobami, kontrolowanie zasobów, udostępnianie usług, administrowanie klientem i dostępem abonenckim, raportowanie o konfiguracji i stanie zasobów.

Jednym z ważnych etapów tworzenia systemu zarządzania ruchem było zapewnienie zgodności jego funkcjonowania z obowiązującymi zaleceniami ITU-T. Wymagało to przeprowadzenia analizy i wyboru metod i parametrów zarządzania ruchem.

3. METODY I PARAMETRY ZARZĄDZANIA RUCHEM

Zagadnienia zarządzania siecią (w tym zarządzania ruchem) są wyspecyfikowane i zdefiniowane w zaleceniach ITU-T serii E.41x [2, 3, 4]. Istotne z punktu widzenia pracy były też zalecenia serii Q.54x dotyczące zarządzania węzłami komutacyjnymi [5].

W zaleceniu E.412 wskazuje się 15 metod zarządzania ruchem. Wybór z tego zbioru konkretnej metody do zastosowania w projektowanym systemie było wynikiem analizy (wspólnej z programistami firmy DGT), w której głównymi kryteriami były: skutki oddziaływania metody na ruch w sieci, łatwość implementacji oraz możliwości systemu komutacyjnego DGT. W rezultacie do zastosowania w systemie zarządzania wybrano następujące metody zarządzania ruchem:

- Tymczasowego kierowania alternatywnego TAR (Temporary Alternative Routing);
- Ukierunkowania wiązki łączy (*Circuit Directionalization*);
- Dziurawienia wywołań (*Call Gapping*);
- Blokowania wiązki łączy (*Circuit Blocking*);
- Anulowania kierowania bezpośredniego w wersjach DRT i DRF (*Cancellation of Direct Routing to or from*);
- Anulowania kierowania alternatywnego w wersjach ART i ARF (*Cancellation of Alternative Routing to or from*);
- Sterowania pomijaniem (*Skip Control*).

Powyższe metody uzupełniono metodą wyznaczania w sieci punktów trudnodostępnych HTR (*Hard-To-Reach*). Wyznaczanie tych punktów ułatwia operatorowi rozpoznanie poziomu obsługi ruchu przez sieć i efektywniejsze stosowanie metod sterowania.

W kolejnym etapie pracy, dla wybranych metody zarządzania określono zestaw niezbędnych parametrów wydajnościowych, których wyznaczenie i obserwacja pozwala na podejmowanie decyzji o uruchomieniu określonej metody zarządzania oraz na ocenę efektywności zastosowanej metody.

Kierując się przede wszystkim tym kryterium, do zastosowania w systemie wybrano następujące parametry:

Współczynnik wywołań obsługowanych ABR (*Answer Bid Ratio*)

$$ABR = \frac{\text{liczba wywołań, które uzyskały sygnał zgłoszenia } Ab \ B}{\text{liczba wszystkich wywołań}} \times 100 \quad (3.1)$$

Parametr ten jest bezpośrednim wskaźnikiem efektywności ruchu. ABR określa relację pomiędzy liczbą wywołań zakończonych stanem komunikacji, a całkowitą liczbą wszystkich wywołań.

Współczynnik zgłoszeń obsługowanych ASR (*Answer Seizure Ratio*)

$$ASR = \frac{\text{liczba zgłoszeń, które uzyskały sygnał zgłoszenia } Ab \ B}{\text{liczba wszystkich zgłoszeń}} \times 100 \quad (3.2)$$

Parametr ten jest bezpośrednim wskaźnikiem efektywności świadczonych usług. ASR określa relację pomiędzy liczbą zgłoszeń zakończonych stanem komunikacji a całkowitą liczbą wszystkich zgłoszeń. Pomiar może być dokonywany zarówno na danej grupie łączy jak i na określonym kierunku.

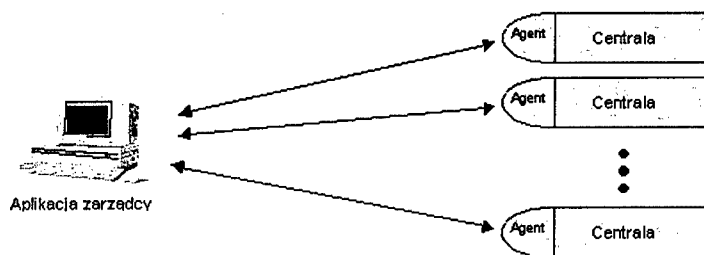
Obciążenie (occupancy)

Obciążenie definiowane jest jako stosunek sumarycznego czasu zajęcia danego obiektu do czasu trwania okresu pomiarowego (czasu obserwacji). Parametr ten może być określany jako całkowite obciążenie dla danej grupy łączy (kierunku) lub jako średnie obciążenie łączy w danej grupie łączy. Parametr ten pozwala identyfikować odbiegające od normy poziomy obsługi ruchu.

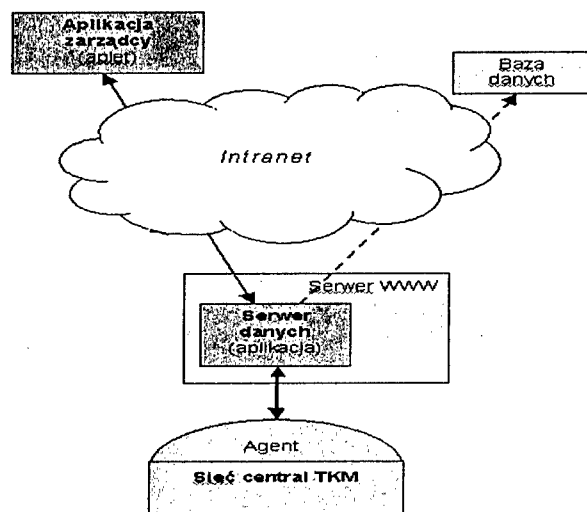
Wybór tych parametrów umożliwi: wyznaczanie w sieci punktów trudnodostępnych HTR, identyfikowanie natłoku na wiązkach łączy, podejmowanie decyzji o zastosowaniu odpowiedniej metody sterowania oraz wnioskowanie o skutkach działania wybranej metody sterowania.

4. ARCHITEKTURA SYSTEMU

System zarządzania ruchem ma architekturę typu zarządca – agent przedstawioną na rys. 1. Opracowana koncepcja funkcjonowania systemu zarządzania która doprowadziła do powstania, przedstawionej na rys. 2, szczegółowej architektury systemu.



Rys. 1. Schemat systemu



Rys. 2. Architektura systemu

Elementy powyższej struktury systemu zarządzania realizują następujące funkcje:

- Aplikacja zarządzcy prezentuje w przejrzystej postaci informacje o stanie sieci i konfiguracji elementów sieci oraz umożliwia operatorowi systemu zarządzania wpływania na stan i konfigurację sieci oraz jej elementów.
- Serwer danych jest centralnym elementem systemu a do jego najważniejszych funkcji należy: przechowywanie informacji o strukturze sieci i stanie jej elementów (liczbie węzłów, numerów punktów sieci sygnalizacji SS7, nazwach węzłów, liczbie grup łączy w sieci oraz o połączeniach pomiędzy węzłami), uaktualnianie informacji o zapisach w tablicach kierowania oraz wartości parametrów konfiguracji w stosowanej metodzie sterowania), zbieranie danych ruchowych z poszczególnych węzłów, wysyłanie poleceń pochodzących od aplikacji zarządzcy, komunikacja z aplikacjami zarządzcy, przesyłanie informacji do aplikacji zarządzcy, przetwarzanie poleceń wydawanych przez aplikację zarządzcy, sprawdzanie uprawnień zarządzcy, rejestracja zmian dokonywanych przez operatora.
- Aplikacja agenta węzła komutacyjnego, której podstawowymi funkcjami są: zbieranie informacji: o stanach ruchowych węzła i o stosowanych w centrali metodach sterowania ruchem, zmiana na polecenie zarządzcy konfiguracji węzła.

Przyjęta architektura systemu zarządzania spełnia następujące podstawowe wymagania:

- podgląd stanu sieci i poszczególnych jej elementów a w szczególności: struktury zarządzanej sieci, listy dostępnych łączy międzywęzłowych, listy aktywnych alarmów w sieci oraz historii alarmów, informacji o wybranym węźle lub wiązce łączy, historii obsługi ruchu w sieci lub węźle,
- modyfikacji konfiguracji elementów sieci a w szczególności modyfikacji tablic kierowania, blokady wiązek (całej wiązki lub części),
- konfigurowania metod kierowania ruchem a w szczególności: uruchamiania metod sterowania, blokowanie kierunków, blokowanie określonego typu ruchu, modyfikowania tablic Call Gappingu,
- monitorowanie stanu sieci i jej elementów,
- generowanie alarmów w wypadkach wystąpienia określonych niekorzystnych stanów w sieci,
- tworzenie raportów o wprowadzanych zmianach w konfiguracji,
- archiwizacja zjawisk zachodzących w sieci.

Jedynym zewnętrznym aktorem opracowanego systemu jest Zarządca komunikuje się z systemem wykonując ciągi interaktywnych operacji, w wyniku których powinien otrzymać satysfakcjonujący go wynik.

Analiza przewidywanych funkcji systemu wraz z przyjętą metodą projektową doprowadziły do opracowania modelu przypadków użycia i modelu klas [1]. Ponieważ zaprojektowany system jest aplikacją w znacznym stopniu rozproszoną, w celu jego pełnego opisu zbiór przypadków użycia poszerzono o te, które zachodzą pomiędzy różnymi częściami funkcjonalnymi systemu zarządzania.

5. WYNIK REALIZACJI PROJEKTU

5.1. Implementacja systemu

Implementacja systemu została wykonana w języku JAVA®. Wybór tego języka został podyktowany niezwykłą elastycznością jego składni. Jest to język w pełni obiektowy, dzięki czemu stworzony przy jego użyciu kod jest w pełni przenośny, łatwy w modyfikacji i daje możliwość stosunkowo łatwej zmiany dowolnej części składowej systemu, przy zastosowaniu się do ustalonych interfejsów pomiędzy obiektami. Również późniejsze zmiany w systemie, a w szczególności jego rozwój staje się bardzo prosty.

Część aplikacyjna została stworzona z wykorzystaniem biblioteki graficznej Swing, natomiast część serwera danych wykorzystuje wyłącznie biblioteki standardowe API.

System został przeznaczony do pracy w systemie Windows, jednakże istnieje możliwość uruchomienia go na platformach Unixowych. Do poprawnego funkcjonowania system wymaga współpracy z siecią Intranet. System został zaprojektowany do pracy w topologii gwiazdy: jeden centralny serwer i wiele komputerów zarządzających. Do komunikacji z bazą danych wykorzystuje się dodatkowe biblioteki sterowników dostępu do bazy danych. W systemie zastosowano wersję bazy danych MySQL, której bezpłatna licencja umożliwia dodawanie, usuwanie i edycję rekordów bazy w wystarczającym zakresie do poprawnego funkcjonowania systemu.

Do poprawnego funkcjonowania wszystkich składników niezbędny jest komputer wyposażony w maszynę wirtualną JAVA® w wersji co najmniej 1.3. Dodatkowo niezbędna jest współpraca serwera danych z serwerem HTTP. Testy systemu przeprowadzane były

w oparciu o bezpłatną wersję serwera HTTP firmy Apache Software Foundation. Zastosowaną wersją był serwer Apache 2.0.45. Do pełnego dostępu zarządcy systemu do bazy danych niezbędne jest posiadanie odpowiedniego oprogramowania do edycji baz danych. W trakcie testów posłużono się programem MySQL-Front w wersji freeware, dostępnej pod adresem <http://www.mysqlfront.de>.

Minimalne wymagania sprzętowe dla aplikacji to: System Windows 98, PII 333MHz, 128MB RAM, 5MB HDD, ekran 800x600. Minimalne wymagania sprzętowe dla serwera danych to: System Windows 2000, PIII 700MHz, 128MB RAM, 100MB HDD.

Aplikacja zarządcy została zbudowana w postaci apletu, uruchamiającego się w oknie odpowiednio skonfigurowanej przeglądarki internetowej. Takie podejście zostało wybrane po to, by zarządcy sieci (domyślnie znajdujący się w różnych lokacjach i często korzystający z różnych komputerów) mogli pracować niezależnie od zmian w systemie. Zastosowanie apletu daje taką przewagę, że cały kod programu pobierany jest poprzez sieć Intranet z jednego serwera. Wadą takiego rozwiązania jest konieczność wyposażenia komputerów zarządcy w odpowiednie oprogramowanie (maszyna wirtualna Javy i przeglądarka WWW). Jednakże jest to zabieg jednorazowy i nie wiąże się z dodatkowymi kosztami, gdyż wymagane oprogramowanie jest udostępniane bezpłatnie i często jest integralną częścią systemu operacyjnego. Wadą zastosowania apletu jest też uzależnienie szybkości uruchamiania aplikacji od zastosowanego łącza intranetowego, gdyż przed jego uruchomieniem, maszyna zarządcy każdorazowo musi pobrać z serwera aktualny kod źródłowy programu.

Serwer danych został zaimplementowany w postaci serwera RMI. Technologia ta umożliwia zdalne komunikowanie się ze sobą różnych fragmentów programu, wykorzystując jako medium komunikacyjne protokół TCP/IP. Zastosowanie RMI pozwala dodatkowo korzystać aplikacji zdalnej bezpośrednio z metod udostępnianych przez serwer RMI, tak jakby zarówno aplikacja jak i serwer działały na w jednym obszarze roboczym lub stanowiły jedną aplikację. Dodatkowo serwer oparty na RMI potrafi obsługiwać jednocześnie wiele aplikacji klientów.

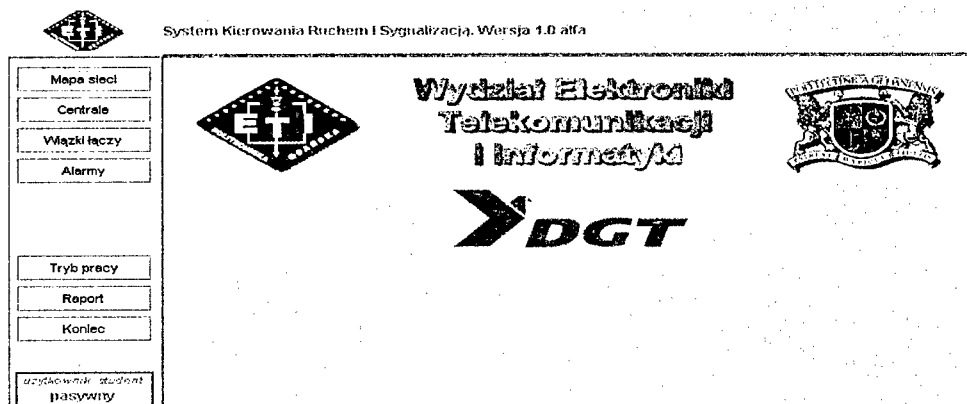
Dodatkowo dla potrzeb testowania systemu a zarazem z myślą zastosowań w laboratorium studenckim system zarządzania został uzupełniony o współpracujący z nim informatyczny model sieci telekomunikacyjnej [1]. Model sieci zaprojektowano tak, aby umożliwić realizację zaimplementowanych metod zarządzania i obserwować ich skutki. Model symuluje ruch w sieci i reaguje na polecenia systemu zarządzania. Model sieci został stworzony w postaci modułu dołączonego do serwera danych. Komunikacja pomiędzy serwerem a modelem sieci odbywa się w oparciu o zaprojektowany protokół komunikacyjny.

5.2. Graficzny interfejs użytkownika

W ramach projektu systemu zarządzania opracowano interfejs graficzny. Okno główne interfejsu, przedstawione na rys. 3, zostało podzielone na części funkcjonalne wizualizacji: struktury sieci, danych o elementach sieci, metod sterowania i elementów sterujących oraz wizualizacja alarmów. Z poziomu tego okna użytkownik ma dostęp do kluczowych funkcji systemu oraz dowiaduje się o aktualnym trybie pracy.

Okno główne podzielone jest na dwie części funkcjonalne. W lewej części okna znajduje się menu, zawierające przyciski umożliwiające wyświetlenie: struktury zarządzanej sieci (Mapa sieci), listy węzłów znajdujących się w sieci (Centrale), listy wiązek łącz w sieci (Grupy łącz), okno aktualnych alarmów oraz historię zarejestrowanych alarmów (Alarmy) oraz pozwala zmienić aktualny tryb pracy, uruchamia funkcję tworzenia raportu

o ruchu w sieci (Raport) a także pozwala bezpiecznie zakończyć pracę z aplikacją. W dolnej części menu znajduje się informacja o aktualnym użytkowniku oraz bieżącym trybie pracy aplikacji.



Rys. 3. Ekran główny aplikacji

5.3. Komunikacja w systemie

System zarządzania komunikuje się z zarządzanymi elementami (węzłami) poprzez aplikacje agentów wykorzystując do tego celu protokół firmowy DGT.

Protokół zarządzania konfiguracją i wymianą danych bazuje na protokole TCP/IP. Występują w nim trzy rodzaje poleceń: polecenia do odczytywania konfiguracji (GET), polecenia do zmiany konfiguracji elementów sieci i przekazywania poleceń sterujących (CMD), polecenia od agenta służące do poinformowania zarządcy o zajściu zdarzenia w centrali (TRAP).

6. TESTY

System, oraz model sieci poddano szeregowi testów sprawdzających poprawność realizacji przyjętych rozwiązań. W szczególności sprawdzono poprawność generowania i przenoszenia zgłoszeń w modelu sieci. Realizując określone scenariusze sprawdzono poszczególne metody sterowania i ich wpływ na parametry wydajnościowe sieci.

System został przetestowany w środowisku Windows™. Testy przeprowadzono na trzech komputerach połączonych w sieć. Zarówno praca systemu, jego stabilność, odporność na nieprzewidziane sytuacje, jak i zaimplementowane mechanizmy sterowania funkcjonowały zgodnie z oczekiwaniami. Przeprowadzone testy poprawności funkcjonowania modelu sieci oraz scenariusze kilku sytuacji wymagających reakcji systemu zarządzania wykazały pełną zgodność z przewidywaniami teoretycznymi i zakończyły się pełnym sukcesem.

7. PODSUMOWANIE

Wynikiem pracy jest kompleksowy system komputerowy umożliwiający zarządzanie siecią w założonym zakresie zaprojektowany zgodnie z wybraną metodą informatyczną.

W systemie zaimplementowano wybrane metody kierowania ruchem oraz konfiguracji elementów sieci. Wybór metod został wykonany po przeprowadzeniu analizy dostępnych metod pod względem kryterium łatwości implementacji i otrzymywanych rezultatów.

Obiektowa natura stworzonego oprogramowania stwarza możliwość kontynuowania prac nad prezentowanym systemem, bez obaw, że ingerencja w kod doprowadzi do nieprzewidzianych sytuacji. Przykładem mogą być wprowadzane pod koniec realizacji projektu nowe funkcjonalności, nie wymagające żadnych zmian w już istniejącym kodzie. Rozwój może dotyczyć zarówno wprowadzania nowych metod sterowania ruchem, próby zaimplementowania automatycznych metod sterowania z uwzględnieniem systemów ekspertowych, przystosowania systemu do pracy z rzeczywistymi obiektami, zmiany protokołu zarządzania na SNMP oraz rozszerzenia systemu o nowe obszary funkcjonalne, np. zarządzanie sygnalizacją.

Możliwe jest wprowadzenie bardziej zaawansowanych metod obróbki i analizy danych otrzymywanych z sieci, co zwiększy efektywność działań operatora przyczyniając się do lepszego wykorzystania istniejących zasobów sieci.

8. BIBLIOGRAFIA

- [1] Stencel A., Świeciak D.R.: Model obiektowy systemu zarządzania ruchem z możliwością zastosowania w centralach DGT. Praca dyplomowa. Politechnika Gdańska 2003
- [2] ITU-T Recommendation E.410 International network management – General information
- [3] ITU-T Recommendation E.411 International network management – Operation guidance
- [4] ITU-T Recommendation E.412 Network management controls
- [5] ITU-T Recommendation Q.543 Digital Exchange Performance Design Objectives
- [6] M. Mycek, R. Weisbrodt: Zarządzanie usługami i sieciami telekomunikacyjnymi – stan i kierunki rozwoju; PTiWT 8-9/2002.
- [7] B. Lauer: *Traffic Network Management – Overview*; wydawnictwo wewnętrzne Lucent Technologies 1998
- [8] *Telecommunications Network Management Into 21st Century*; praca zbiorowa pod redakcją S. Aidarous'a i T. Plevyaka; IEEE Press New York 1994.

TRAFFIC ENGINEERING MANAGEMENT SYSTEM IN NETWORK WITH DGT EXCHANGES

Summary

This paper presents informatics system designed in cooperation with DGT company intended for traffic engineering management using selected methods in general telephone network with DGT exchanges. This paper consists of two parts. The first part deals with results of traffic engineering management methods analysis regarding possibility of using these methods in the network with DGT exchanges. Necessary parameters for selected management methods specified during monitoring of network nodes are included here. The second part presents system requirements and functionality specifications defined method and how to accomplish them. It was described process of software creating in compliance with used methodology from the idea to the end product with fully functionality. Also results of performed tests are discussed.

Wojciech Molisz

Katedra Systemów Informacyjnych, Politechnika Gdańska

MODEL PRZEŻYWALNYCH SIECI IP NASTĘPNEJ GENERACJI

Streszczenie

W pracy rozważa się przeżywalne sieci o warstwowej architekturze logicznej IP-MPLS/ WDM. Węzły analizowanej sieci mają zintegrowaną funkcjonalność ruterów IP-MPLS w warstwie IP oraz przełącznic optycznych (OXC) w warstwie transmisji optycznej. Przepustowości łączy mogą być tak duże (nawet kilkaset Gbit/s), że awaria nawet pojedynczego łącza może dotknąć bardzo liczne grupy użytkowników. Dlatego zasadnicze znaczenie ma zapewnienie przeżywalności tych sieci.

W pracy rozważa się zadania optymalizacji strumieni informacji (IP), jakie sieć ma przesłać w warstwie optycznej między zadanymi parami użytkowników przy możliwie najmniejszym koszcie, spełnieniu wszystkich standardowych ograniczeń w taki sposób, aby uszkodzenie pojedynczego węzła lub światłowodu nie przeszkodziło w transmisji. Przedstawiono model umożliwiający badanie przeżywalnych sieci IP oraz przykładowe wyniki symulacji i wnioski.

1. WSTĘP

W pracy rozważa się sieci o architekturze logicznej, w której w warstwie sieciowej występuje protokół IP zaimplementowany na optycznej sieci transportowej bez pośrednictwa usług i protokołów sieci synchronicznej (SDH). Zdaniem wielu autorów, do takiej architektury migruje sieć następnej generacji w warunkach konwergencji informatyki, mediów i telekomunikacji [1], [4], [10], [11].

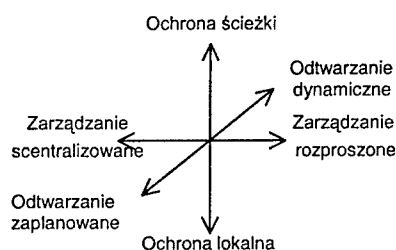
Węzły analizowanej sieci mają zintegrowaną funkcjonalność ruterów IP w warstwie IP oraz przełącznic optycznych (OXC) w warstwie transmisji optycznej. W sieci optycznej stosowane jest zwielokrotnienie za pomocą wielu długości fal (WDM). Dla spełnienia rosnących wymagań użytkowników niezbędne jest zapewnienie przewidywalnej jakości usług IP. Zakłada się, że może to zapewnić rodzina protokołów komutacji według etykiet (MPLS). Zasadnicze znaczenie dla zaakceptowania takich sieci przez użytkowników ma zapewnienie ich przeżywalności.

Zadanie optymalizacji jest następujące: należy zapewnić przesłanie określonych strumieni informacji (IP) w warstwie optycznej przy możliwie najmniejszym koszcie, spełnieniu wszystkich standardowych ograniczeń (przepustowości, liczby długości fal, itp.) w taki sposób, aby uszkodzenie pojedynczego węzła lub światłowodu nie przeszkodziło w transmisji. Omówiono metody rozwiązania zadania optymalizacji zabezpieczenia przed awariami pojedynczych elementów sieci oraz przykładowe wyniki symulacji.

2. METODY ZAPEWNIENIA PRZEŻYwalNOŚCI ARCHITEKTURY WARSTWOWEJ

2.1. Metody zapewnienia przeżywalności sieci o topologii kratowej w jednej warstwie

Klasyfikację metod zapewnienia przeżywalności sieci o topologii kratowej opieramy na trzech podstawowych kryteriach, przedstawionych na rys. 1 [7].



Rys. 1. Klasyfikacja metod zapewnienia przeżywalności sieci o topologii kratowej

Zaletą scentralizowanego zarządzania procesami naprawy sieci przeżywalnych o topologii kratowej jest globalnie optymalny rozdział zasobów sieci, a główną wadą – długi czas wyznaczania nowej konfiguracji sieci. Czas ten jest znacznie krótszy dla zarządzania rozproszonego, którego zaletą jest też większa odporność sieci na awarie, natomiast wadą – mniej efektywny rozdział zasobów sieci.

Drugim kryterium klasyfikacyjnym jest relacja chwili określenia planu rekonfiguracji sieci do chwili wystąpienia awarii. Jeżeli plan ten jest przygotowany przed wystąpieniem awarii to mówimy o zaplanowanej rekonfiguracji. Jej zaletą jest krótki czas rekonfiguracji sieci w porównaniu z czasem trwania rekonfiguracji dynamicznej, w której poszukiwanie nowych tras obejściowych rozpoczyna się dopiero po wystąpieniu awarii, natomiast wadą – konieczność rezerwacji większych zasobów sieci.

Trzecim kryterium klasyfikacyjnym na rys. 1 jest rodzaj chronionych zasobów: całej ścieżki od krańca do krańca (ang. *end-to-end*) lub tylko pojedynczego łącza (ochrona lokalna). Warto podkreślić, że ścieżka zabezpieczająca od krańca do krańca nie musi być dłuższa od ścieżki podstawowej; w sieci z ochroną lokalną, ścieżka obejściowa jest zawsze dłuższa od ścieżki podstawowej.

Ze względu na wzajemną rozłączność przedstawionych kryteriów klasyfikacyjnych wchodzą w grę różne metody zapewnienia przeżywalności a więc i różne metody odtwarzania po awariach.

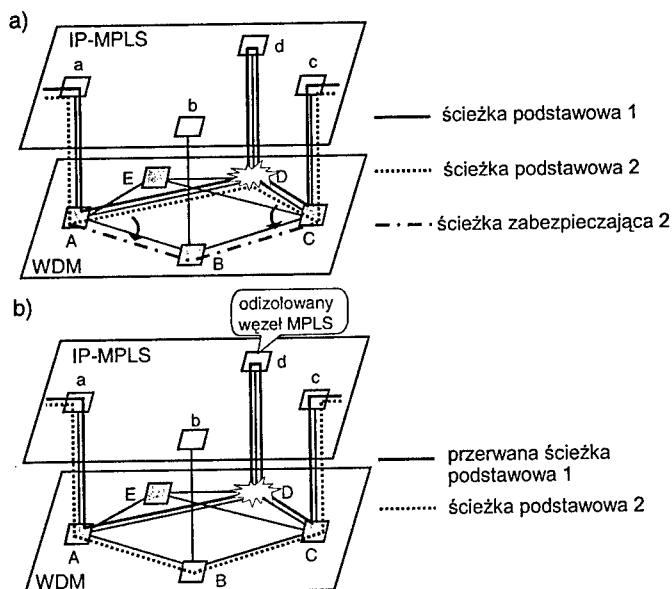
2.2. Metody zapewnienia przeżywalności warstwowej architektury sieci

Różne mechanizmy zapewnienia przeżywalności sieci występują z reguły w kilku warstwach architektury logicznej. Podstawowe znaczenie ma decyzja, która warstwa lub które warstwy są odpowiedzialne za zapewnienie przeżywalności sieci.

Zajmiemy się odtwarzaniem zaplanowanym.

2.2.1. Przeżywalność zapewniana przez warstwę dolną (WDM)

Odtwarzanie strumieni dotkniętych awarią węzła D, przeprowadzane w dolnej warstwie architektury logicznej sieci (WDM) jest zilustrowane na rys. 2.



Rys. 2. Ilustracja działania mechanizmów odtworzeniowych w dolnej warstwie:
a) stan po awarii; b) stan po rekonfiguracji sieci

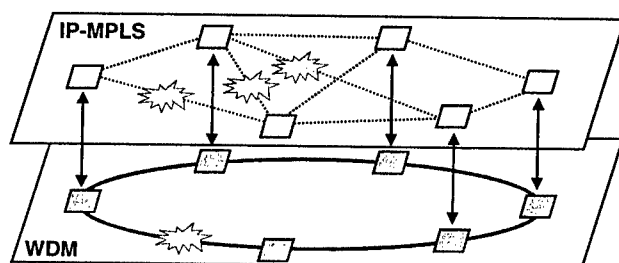
Zaletą zapewnienia przeżywalności sieci przez dolną warstwę jest mała liczba działań niezbędna do zapewnienia obejścia uszkodzonego elementu sieci oraz ograniczenie działań tylko do jednej warstwy. Wadą jest brak możliwości przywracania połączeń w każdej warstwie powyżej tej, w której nastąpiło uszkodzenie.

Przykładowo, po awarii węzła D w warstwie WDM (rys. 2) możliwe jest tylko przekierowanie tych strumieni danych, które przepływały przez ten węzeł w dolnej warstwie (ścieżka podstawowa 2). Węzeł d warstwy wyższej (IP-MPLS) pozostaje izolowany i strumień, który był przez niego przesyłany (ścieżka podstawowa 1) nie może być odtworzony przez warstwę WDM.

2.2.2. Przeżywalność zapewniana przez warstwę górną (IP-MPLS)

Alternatywą przedstawionego podejścia jest powierzenie odpowiedzialności za przeżywalność warstwie górnej architektury logicznej sieci (IP-MPLS). Złożoność działań naprawczych jest teraz większa, ponieważ jedno uszkodzone połączenie warstwy dolnej może pociągać dużą liczbę uszkodzonych połączeń nawiązanych przez warstwę górną, jak zilustrowano na rysunku 3.

Zalet odtwarzania strumieni przez warstwę górną jest kilka. Po pierwsze można w pierwszej kolejności odtwarzać strumień o najwyższych priorytetach.



Rys. 3. Pojedyncze uszkodzenie w warstwie WDM powoduje wiele uszkodzeń w warstwie IP-MPLS

Po drugie możliwe jest rozdzielanie przywracanych połączeń pomiędzy wiele alternatywnych ścieżek. Po trzecie, można bardziej równomiernie wykorzystać zasoby sieci przez przekierowanie strumieni z uszkodzonych ścieżek na ścieżki z dużymi rezerwami przepustowości. Nie jest to możliwe do przeprowadzenia przez warstwę dolną, ponieważ ta przełącza wszystkie strumienie jednocześnie.

2.2.3. Przeżywalność zapewniana przez wszystkie warstwy

W przypadku współodpowiedzialności warstw za przeżywalność można wyróżnić dwa główne podejścia. W podejściu „oddolnym” (ang. *bottom-up*) działania naprawcze są podejmowane w dolnej warstwie. Dopiero strumienie, których nie można odtworzyć przez tę warstwę spróbuje przywrócić warstwa wyższa. Zaletą tego podejścia jest ograniczenie liczby analizowanych elementów sieci.

Z kolei w podejściu „odgórnym” (ang. *top-down*) działania naprawcze podejmowane są najpierw w warstwie górnej, a dopiero potem (w razie niepowodzenia) – w warstwie niższej. Zaletą tego podejścia jest możliwość uwzględnienia priorytetów ruchu, a wadą – konieczność dodatkowej sygnalizacji pozwalającej stwierdzić czy warstwa wyższa była w stanie odtworzyć strumienie czy też nie.

W dalszym ciągu zakładamy ochronę ścieżki, odtwarzanie zaplanowane, zarządzanie rozproszone, naprawę „odgórną”.

3. MODEL SIECI PRZEŻYWALNEJ

3.1. Zadanie optymalizacji rozplywu strumieni informacji i rozdziału długości fal

Zadanie optymalizacji rozplywu strumieni IP jest następujące. Dane są natężenia strumieni informacji (IP), jakie sieć ma przesłać między zadanymi (lub wszystkimi) parami użytkowników. Należy zapewnić przesłanie tych strumieni w warstwie IP przy możliwie najmniejszym koszcie, spełnieniu wszystkich standardowych ograniczeń (przepustowości, nieujemności itp.) w taki sposób, aby uszkodzenie pojedynczego węzła lub łącza (światłowodu) nie zakłóciło transmisji. Wymagania przeżywalności można spełnić dwójako: planując dedykowane ścieżki zabezpieczające dla ścieżek podstawowych przenoszących żądane strumienie lub planując ścieżki współdzielone (jedna ścieżka zabezpiecza więcej niż jedną ścieżkę podstawową). Zakładamy, że każda ścieżka zabezpieczająca jest węzłowo

rozłączna ze swoją ścieżką podstawową, aby zapobiec jednoczesnej awarii obu ścieżek wykorzystujących różne długości fal w tym samym światłowodzie [12].

Zadanie optymalizacji można formalnie zapisać jako mieszane zadanie optymalizacji liniowej całkowitoliczbowej (zadanie optymalizacji przepływów wieloskładnikowych). Zadania tego typu są rozważane między innymi w [6], [10] - [12]. Niestety zmienne całkowite, wprowadzone przez warunki rozłączności węzłowej tras zabezpieczających i podstawowych powodują, że zadanie jest NP-zupełne [3].

Następnie optymalne strumienie IP należy odwzorować na ścieżki optyczne w warstwie WDM z uwzględnieniem ograniczonej liczby długości fal. Te ograniczenia powodują, że zadanie optymalnego odwzorowania strumieni IP na ścieżki optyczne jest też NP-zupełne [3]. Z tych względów, a także z powodu braku efektywnych narzędzi (takich, jak np. program CPLEX firmy ILOG), zadania optymalizacji rozwiążemy za pomocą algorytmów heurystycznych. Z powodu ograniczonej objętości referatu omówimy tylko przypadek dedykowanych ścieżek zabezpieczających.

3.2. Metoda wyznaczenia dedykowanych ścieżek podstawowych i zabezpieczających

Wyznaczenie ścieżki podstawowej obejmuje kolejne kroki:

1. Wyznaczenie K najkrótszych ścieżek łączących węzeł źródłowy z docelowym, np. w oparciu o algorytm opisany w [2].
2. Użycie najkrótszej ścieżki ze zbioru ścieżek znalezionych w kroku 1.

Wyznaczenie ścieżki zabezpieczającej:

3. Wyznaczenie najkrótszej ścieżki łączącej węzeł źródłowy z docelowym przy uwzględnieniu węzłowej rozłączności ścieżki podstawowej oraz zabezpieczającej.
4. Jeśli ścieżka taka zostanie znaleziona, jest przydzielona strumieniowi; w przeciwnym przypadku przydzielona ścieżka podstawowa jest usuwana, analizowana jest kolejna z K najkrótszych ścieżek podstawowych znalezionych w kroku 1 i ponownie podejmowana jest próba znalezienia ścieżki zabezpieczającej. Jeżeli nie da się znaleźć ścieżki zabezpieczającej dla jakiegokolwiek ze znalezionych ścieżek podstawowych sygnalizowana jest niemożność zestawienia obu ścieżek dla zadanego strumienia.

3.3. Metoda odtwarzania strumieni ruchu uszkodzonych na skutek awarii sieci

W pierwszym etapie kolejno sprawdzane są wszystkie strumienie ruchu, którym zostały przydzielone ścieżki podstawowe i ścieżki zabezpieczające. Jeżeli strumień podstawowy został uszkodzony, podejmowana jest próba przeniesienia go na ścieżkę zabezpieczającą. To może się nie udać, jeżeli ścieżka zabezpieczająca również uległa uszkodzeniu.

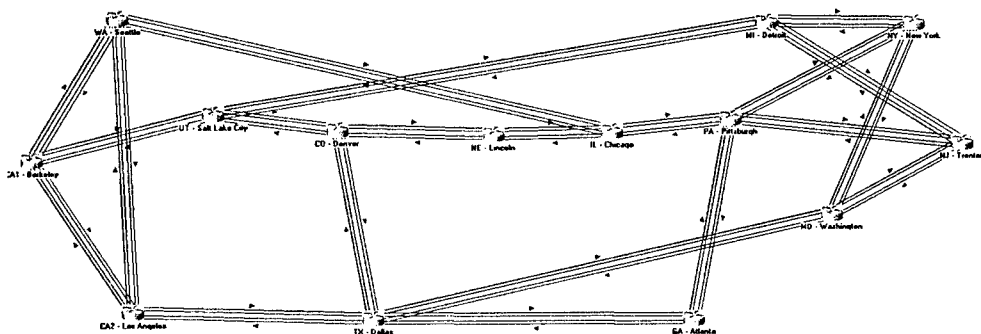
W drugim etapie podejmowane są próby odtworzenia tych strumieni ruchu, których naprawa nie powiodła się. Kolejno sprawdzane są wszystkie strumienie ruchu i podejmowane są następujące działania:

1. Dla strumieni, które w poprzednim etapie nie zostały przeniesione na swoją ścieżkę zabezpieczającą z powodu jej uszkodzenia, podejmowana jest próba wyznaczenia zarówno nowej ścieżki podstawowej, jak i zabezpieczającej.
2. Dla strumieni, które w poprzednim etapie zostały przeniesione na swoją ścieżkę zabezpieczającą podejmowana jest próba wyznaczenia nowej ścieżki zabezpieczającej.

W ostatnim, trzecim etapie próbuje się znaleźć nowe ścieżki zabezpieczające dla strumieni, w których właśnie ścieżki zabezpieczające zostały uszkodzone.

4. PRZYKŁADOWE WYNIKI MODELOWANIA

W celu sprawdzenia poprawności omawianego modelu oraz zbadania efektywności algorytmów została przeprowadzona symulacja różnych sieci. Przedstawimy tutaj wyniki badania sieci NSF, pokazanej na rys. 4. Program napisał i obliczenia wykonał Muzi ot [8].



Rys. 4. Topologia modelowanej sieci NSF

W sieci zostały wyznaczone strumienie ruchu, rozpoczynające się i kończące w losowo wybranych węzłach.

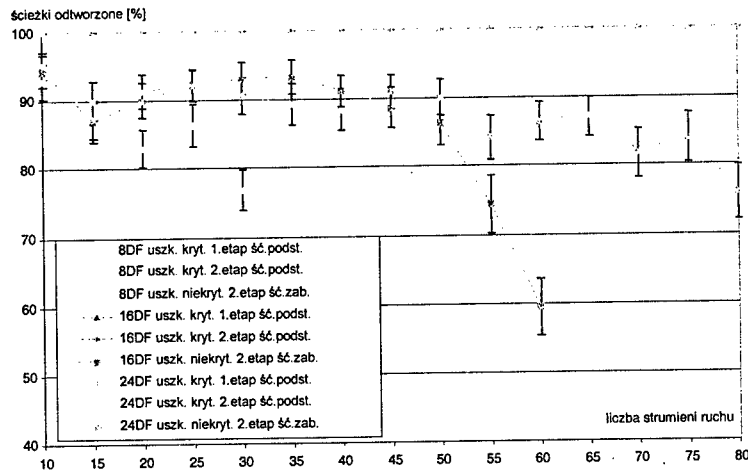
Na rys. 5 przedstawiono wykresy procentowego odtworzenia strumieni ruchu w funkcji liczby zdefiniowanych strumieni ruchu dla sieci obsługujących 8, 16 i 24 długości fal dla uszkodzenia pojedynczego światłowodu, a na rys. 6 – dla uszkodzenia węzła. Awaryjną ścieżkę podstawową nazwano krytyczną, a ścieżki zabezpieczające – niekrytyczną. Na rysunkach zaznaczono przedziały ufności dla 95% poziomu ufności.

Jak widać na rys. 5, po uszkodzeniu pojedynczego światłowodu dane z przerwanych ścieżek podstawowych są w 100% przenoszone na ścieżki zabezpieczające. Następnie podejmowane są próby wyznaczenia nowych ścieżek zabezpieczających. Niestety nie udaje się wyznaczyć nowych ścieżek zabezpieczających dla wszystkich strumieni. W miarę wzrostu liczby strumieni ruchu spada współczynnik odtworzonych ścieżek zabezpieczających.

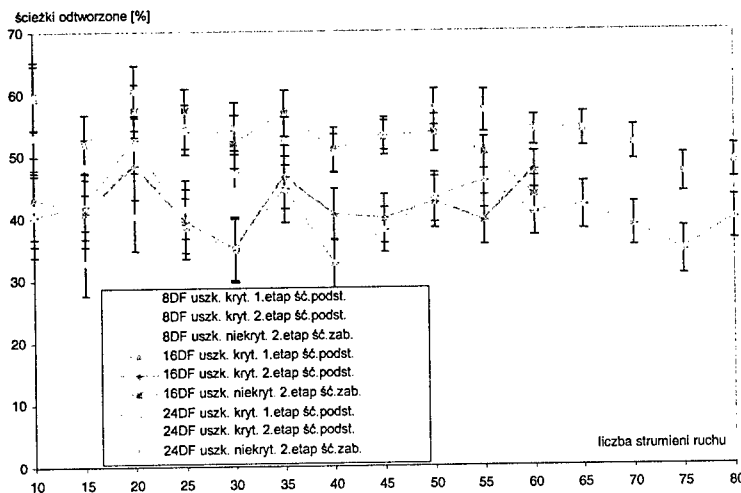
Kolejny typ uszkodzenia – awaria pojedynczego węzła przynosi już znaczący spadek współczynnika odtwarzania strumieni ruchu. Jak widać na rys. 6, w pierwszym etapie udaje się odtworzyć w przybliżeniu od 35% do 50% ścieżek z zabezpieczeniem dedykowanym.

5. PODSUMOWANIE

Przedmiotem rozważań niniejszej pracy były przeżywalne sieci IP-MPLS/WDM. Opisano metody zabezpieczania przeżywalności sieci i przywracania jej możliwie pełnej funkcjonalności. Zadania optymalizacji doboru tras dla strumieni ruchu IP i przydzielania im odpowiednich długości fal zostały rozwiązane za pomocą algorytmów heurystycznych.



Rys. 5. Rodzina wykresów procentowego odtworzenia strumieni ruchu w funkcji liczby zdefiniowanych strumieni ruchu dla sieci NSF obsługujących 8, 16 i 24 długości fal po uszkodzeniu pojedynczego światłowodu.



Rys. 6. Rodzina wykresów procentowego odtworzenia strumieni ruchu w funkcji liczby zdefiniowanych strumieni ruchu dla sieci NSF obsługujących 8, 16 i 24 długości fal po uszkodzeniu pojedynczego węzła.

Dokładność zastosowanych algorytmów heurystycznych została zweryfikowana w [8] przez porównanie z wynikami [9]. Zgodność otrzymanych wyników okazała się dobra.

Przykładowa symulacja sieci NSF pokazała, że przy uszkodzeniu pojedynczego światłowodu odtworzonych zostaje 100% strumieni z uszkodzonych ścieżek podstawowych. Ponadto odtworzonych zostało 80 do 90%, uszkodzonych ścieżek zabezpieczających. W przypadku uszkodzenia węzłów, wartości procentowego odtworzenia strumieni ruchu znacznie spadały. Wynika to z faktu, że znaczny odsetek nie odtworzonych strumieni

ruchu rozpoczynał się lub kończył w uszkodzonym węźle. Otrzymane wyniki potwierdziły słuszność przyjętych założeń modelu i pozwoliły na zbadanie ważnych aspektów funkcjonowania sieci IP następnej generacji.

BIBLIOGRAFIA

- [1] *After the Optical Bubble: The Reality Check* (7 artykułów), IEEE Communications Magazine, Sept. 2003.
- [2] Bhandari R.: *Survivable Networks: Algorithms for Diverse Routing*, Kluwer, Boston 1999.
- [3] Chlamtac I., Ganz A., Karmi G.: *Lightpath communications: An approach to high-bandwidth optical WAN's*, IEEE Trans. Commun., vol. 40, pp. 1171–1182, July 1992.
- [4] Colle D., De Maesschalck S., Develder C., Van Heuven P., Groebbens A., Cheyns J., Lievens I., Pickavet M., Lagasse L., Demeester P.: *Data-Centric Optical Networks and Their Survivability*, IEEE J. Select. Areas Commun., Vol. 20, pp. 6–21, 2002.
- [5] Crochat O., Le Boudec J-Y., Gerstel O.: *Protection Interoperability for WDM Optical Networks*, IEEE/ ACM Trans. Networking, Vol. 8., pp. 384-395, 2000.
- [6] Doshi B.T., Dravida S., Harshavardhana P., Hauser O., Wang Y.: *Optical network design and restoration*, Bell Labs Technical Journal, pp. 58–83, Jan.-March 1999.
- [7] Molisz W.: *Przeżywalność sieci teleinformatycznych i telekomunikacyjnych*, Politechnika Gdańska, Monografie Nr. 31, Gdańsk, 2002.
- [8] Muzioł M.: *Samonaprawialne sieci MPLS*, Praca magisterska, promotor: W. Molisz, Politechnika Gdanska, Wydział ETI, Gdańsk 2003.
- [9] Ramamurthy S., Mukherjee B.: *Survivable WDM Mesh Networks, Part II – Restoration*, Proc. IEEE ICC'99, vol. 3, Vancouver, BC, pp. 2023–2030, June 1999.
- [10] Sahasrabuddhe L., Ramamurthy S., Mukherjee B.: *Fault Management in IP-Over-WDM Networks: WDM Protection Versus IP Restoration*, IEEE J. Select. Areas Commun., Vol. 20, pp. 21–33, 2002.
- [11] Sridharan M., Somani A. K., and Salapaka M. V.: *Approaches for capacity and revenue optimization in survivable WDM networks*, J. High Speed Networks, vol. 10, no. 2, pp. 109–125, 2001.
- [12] Zang H., Ou C., Mukherjee B.: *Path-Protection Routing and Wavelength Assignment (RWA) in WDM Mesh Networks Under Duct-Layer Constraints*, IEEE/ACM Trans. Networking, Vol. 11., pp. 248-258, 2003.

SURVIVABLE NEXT GENERATION IP NETWORKS

Summary

This paper deals with survivable networks having layered architecture IP-MPLS/WDM. Nodes have integrated functionality of IP-MPLS routers in the IP layer and of optical cross connects in the optical transmission layer. Since link capacities may approach several hundreds of gigabits per second, even a single link failure may affect very large groups of users. Network survivability becomes really the fundamental issue.

The IP flow optimization problem, satisfying given demands at minimum cost and protecting flows in the optical layer against a single arc (or a single node) failure, is considered. Methods solving the problem and simulation results are described.

Tomasz Mrugalski

Katedra Systemów Informacyjnych, Politechnika Gdańska

METODA SZYBKIEGO WDRAŻANIA IPV6 PRZY WYKORZYSTANIU DHCPV6

Streszczenie

Podstawowe założenia i wymagania dla protokołów IP nowej generacji zostały opracowane już w roku 1995. Pomimo ich ogromnych możliwości, rozwiązania sprzętowe i programowe wykorzystujące IPv6 pojawiają się wolniej niż powszechnie oczekiwano. W skali sieci lokalnych głównym problemem okazuje się brak ugruntowanej struktury usługodawców, a także złożony proces konfiguracji. Niniejsze opracowanie analizuje poszczególne fazy konieczne do osiągnięcia funkcjonalności IPv6 w sieci LAN – metody tunelowania, problemy związane z konfiguracją routera oraz hostów. Przedstawiona została propozycja metody szybkiego wdrożenia IPv6 w sieci lokalnej przy użyciu rozwiązań opartych o implementację protokołów DHCPv6 oraz niektórych aspektów ICMPv6.

1. WSTĘP

Rodzina protokołów IPv6 [1] jest następcą, będącej obecnie w powszechnym użyciu, rodziny protokołów IP w wersji 4. Liczy sobie ona już ponad 20 lat. Protokół ten nie był nigdy projektowany z myślą o sieci o zasięgu globalnym, a jedynie do komunikacji między ośrodkami naukowymi na terenie Stanów Zjednoczonych. Te założenia wiązały się z określonymi cechami protokołu, np. dużą wagę przykładano do sprawdzania poprawności przesyłanych danych, natomiast w ogóle nie przewidziano ochrony kryptograficznej. Na przestrzeni dwóch dekad, opracowano wiele opcjonalnych rozwiązań, nierzadko wzajemnie niekompatybilnych, pozwalających uniknąć lub minimalizować pewne konkretne wady.

Pojawił się również inny problem, który wymaga radykalnego podejścia - kończąca się przestrzeń adresowa. Istniejące rozwiązania pośrednie – translacja adresów (*NAT*) lub routing bezklasowy (*CIDR*) – nie są w stanie sprostać wymaganiom. Translacja adresów w konfiguracji wiele do jeden, czyli tzw. *maskarada*, umożliwia tylko jednokierunkowe zestawienie połączenia, co uniemożliwia poprawne działanie wielu mechanizmów, a więc w efekcie powoduje, że liczne protokoły stają się bezużyteczne. Konieczne jest więc rozszerzenie przestrzeni adresowej. To oferuje właśnie IPv6.

1.1. Wdrażanie IPv6 -- fazy

Wdrażenie IPv6 w sieci lokalnej można podzielić na 3 fazy: uzyskanie puli adresowej, konfigurację routera oraz konfigurację hostów. Każda z tych faz została poniżej krótko scharakteryzowana.

1.2. Faza 1 -- uzyskanie puli adresowej

W obecnym Internecie ruch pakietów IPv6 należy wciąż do rzadkości. Dlatego też konieczne jest wykorzystanie tzw. tuneli, czyli przenoszenia danych jednego protokołu wewnątrz drugiego. W tym przypadku ruch IPv6 jest przenoszony za pomocą klasycznego IPv4. Cała procedura nazywana jest popularnie *tunelowaniem*. W chwili obecnej zestawienie tunelu jest najprostszą i w wielu przypadkach jedyną drogą do uzyskania puli adresowej IPv6. Obecnie wiele ośrodków umożliwia darmowe zestawianie tuneli, a także przydziela do nich pule adresowe. Procedura rejestracji, zestawienia oraz uruchomienia tunelu zajmuje przeciętnemu administratorowi około 15 minut.

1.3. Faza 2 -- konfiguracja routera

Po zakończeniu konfiguracji tunelu, dostęp do IPv6 możliwy jest wyłącznie z routera. Należy wykonać pewne kroki, aby możliwa była obsługa hostów znajdujących się w sieci LAN. Konieczny jest podział posiadanej klasy na mniejsze, skonfigurowanie tras oraz zadbanie o obsługę hostów. W tym celu niezbędne jest rozgłaszanie informacji o dostępnych trasach, a także uruchomienie serwera DHCPv6 [2].

1.4. Faza 3 -- konfiguracja hostów

Ostatnią fazą jest konfiguracja hostów. Z punktu widzenia całej operacji niezwykle istotne jest, aby była ona możliwie najkrótsza, ponieważ będzie ona powtarzana na każdym hoście znajdującym się w sieci. Na szczęście dzięki założeniom projektowym klienta DHCPv6, a także obowiązkowej obsłudze mechanizmu Router Advertisement ([3], [4]), zadanie to sprowadza się do instalacji klienta DHCPv6, co w przeciętnych warunkach zajmuje około 5 minut.

2. PULA ADRESOWA

Podział całej przestrzeni adresowej został zrealizowany zgodnie z zaleceniami i wytycznymi *Internet Engineering Task Force*. Struktura administracyjna przestrzeni adresowej jest ściśle hierarchiczna. Dla przykładu w Europie przedzielaniem adresów zajmuje się organizacja *RIPE*. Delegowała ona podsieć 2000:0a10::/32 do dyspozycji *Naukowo-Akademickiej Sieci Komputerowej*, która zajmuje się przydzielaniem pul adresowych na terenie Polski. NASK delegował mniejsze pule adresowe (np. /48) operatorom regionalnym. Podobne założenia sprawdzają się w przypadku zastosowania routingu bezklasowego (CIDR) dla puli adresów IPv4.

Problemem jest jednak przenoszenie ruchu IPv6. W tej chwili znakomita większość routerów nie obsługuje tego typu ruchu, dlatego nie jest możliwe jego kierowanie bezpośrednio. Trzeba w tym celu posłużyć się mechanizmem tzw. tuneli.

2.1. Koncepcja tuneli

W chwili obecnej tylko nieliczne sieci obsługują ruch IPv6. Bezpośrednia transmisja pomiędzy nimi nie jest zazwyczaj możliwa. Taki stan rzeczy określany jest się często mianem „wysp” IPv6. Połączenie pomiędzy poszczególnymi wyspami jest realizowane za pomocą przenoszenia ruchu IPv6 z wykorzystaniem innego protokołu. W ogromnej większości przypadków jest to klasyczny IPv4. Proces przenoszenia jednego protokołu za pomocą innego nazywamy enkapsulacją, a procedurę łączenia nazywamy tunelowaniem.

2.2. Tunelowanie

Istnieje kilka technik tunelowania. Do bardziej znanych należą: *ip-ip*, *Generic Routing Encapsulation*, oraz *Simple Internet Transition*. Technika *ip-ip* służyła do tunelowania IPv4 po IPv4 i była używana głównie w środowisku linuksowym. Obecnie nie jest już używana i została zastąpiona nowszymi metodami: *GRE* oraz *SIT*. *GRE* jest uniwersalnym protokołem tunelowania wprowadzonym przez firmę *Cisco*. Umożliwia ona tunelowanie dowolnego protokołu za pomocą IPv4. Jest on dziś powszechnie stosowany w wielu różnych zastosowaniach. Powstała również metoda tunelowania przeznaczona specjalnie do przenoszenia ruchu IPv6 po IPv4 – *Simple Internet Transition*. Jest ona najczęściej wykorzystywaną metodą w zastosowaniach IPv6. W prezentowanym poniżej przykładzie została użyta właśnie ta metoda.

2.3. Opis rozwiązania przykładowego

Procedura uzyskania tunelu rozpoczyna się od znalezienia odpowiedniego usługodawcy. Omawiane tu rozwiązanie wykorzystuje tunel udostępniony nieodpłatnie przez providera internetowego „Hurricane Electric”. Po rejestracji na stronie [www \(http://tunnelbroker.net\)](http://tunnelbroker.net) uzyskujemy dostęp do panelu konfiguracyjnego. Kluczowym parametrem, który można określić, jest adres IPv4 lokalnego końca tunelu. W przypadku laboratorium katedry KSI wydziału ETI Politechniki Gdańskiej jest to 153.19.48.120. Niezbędna jest konfiguracja lokalna. W tym celu należy przygotować kernel linuksowy z obsługą IPv4, IPv6, tunelowania typu *SIT* oraz przekazywania pakietów. Minimalny skrypt konfiguracyjny tunel został przedstawiony poniżej.

```
ip tunnel add tunel0 mode sit local $LOCAL remote $REMOTE ttl
64 dev eth0
ip link set tunel0 up
```

Po wykonaniu powyższych czynności, dysponujemy zestawionym tunelem. Przez system operacyjny jest on widoczny jako interfejs *tunel0*. Nie jest on jednak jeszcze skonfigurowany do przesyłania danych IPv6. W tym celu należy przypisać mu odpowiedni adres:

```
ip addr add 2001:470:1f01:ffff::2b3/127 dev tunel0
ip route add 2000::/3 dev tunel0
```

Należy zwrócić uwagę na czynności wykonane przez system po wydaniu tego polecenia. Oprócz przypisania adresu do interfejsu, została również skonfigurowana trasa do jednobitowej klasy, do której należą adresy 2001:470:1f01:ffff::2b2 oraz

2001:470:1f01:ffff::2b3. Są to zdalny i lokalny adres IPv6 końców tunelu. Po wykonaniu tych czynności możliwe jest zrealizowanie połączenia na poziomie IPv6. Ostatnią czynnością jest określenie domyślnej trasy dla pakietów IPv6. Dane będą przesyłane przez stworzony właśnie tunel.

Dzięki wykonaniu tej ostatniej czynności proces konfiguracji tunelu został w pełni zakończony, a serwer dysponuje pełnym dostępem do usług opartych o protokół IPv6. Przykładowy rezultat działania polecenia `ping6 www.kame.net -n` został przedstawiony poniżej:

```
PING www.kame.net(2001:200:0:8002:203:47ff:fea5:3085) from
3ffe:8320:3:2ff::6 : 56 data bytes
64 bytes from 2001:200:0:8002:203:47ff:fea5:3085: icmp_seq=1
ttl=55 time=453 ms
64 bytes from 2001:200:0:8002:203:47ff:fea5:3085: icmp_seq=2
ttl=55 time=649 ms
64 bytes from 2001:200:0:8002:203:47ff:fea5:3085: icmp_seq=3
ttl=55 time=585 ms
--- www.kame.net ping statistics ---
3 packets transmitted, 3 received, 0% loss, time 2017ms
rtt min/avg/max/mdev = 453.606/562.669/649.364/81.467 ms
```

3. KONFIGURACJA ROUTERA

W typowym przypadku funkcje routera oraz serwera pełni ten sam węzeł¹. W prezentowanym tu rozwiązaniu jest to system linuksowy. Procedura jego konfiguracji składa się z czterech części – statycznej konfiguracji routingu, uruchomienia demona routingu, uruchomienia usługi rozwiązywania nazw oraz uruchomienia serwera DHCPv6.

3.1. Routing

Konieczna jest obsługa zarówno ruchu przychodzącego z Internetu, jak i wychodzącego z sieci lokalnej. Konfiguracja domyślnej trasy (2000::/3) została przeprowadzona przy okazji konfigurowania tunelu. Pozostaje jeszcze przydzielić odpowiednią pulę adresową sieci lokalnej. Niezbędne jest również włączenie przekazywania pakietów.

```
ip -6 route add 2001:470:1f01::/80 dev eth1
cat 1 >/proc/sys/net/ipv6/conf/all/forwarding
```

W tym momencie hosty z sieci mogą już wysyłać swoje pakiety IPv6. Wymagana jest jednak ręczna konfiguracja zarówno własnego adresu, domyślnego routera, jak i serwera DNS, co w przypadku większej ilości hostów jest nie do przyjęcia.

¹ ang. node. Zgodnie z terminologią dokumentów RFC urządzenie obsługujące protokół IPv6 określane jest mianem węzła. Węzły dzielone są na przekazujące ruch nieadresowany do nich (routery) oraz nieprzekazujące takiego ruchu (hosty).

3.2. Demon routingu

Problem domyślnego routera rozwiązuje uruchomienie demona protokołu *Router Renumbering*. W przypadku środowisk linuksowych demon ten nosi nazwę *radvd*. Okresowo rozsyła on za pomocą wiadomości ICMPv6 typu 134, informacje o dostępnych trasach wraz z ich parametrami.

W prezentowanym tu rozwiązaniu, główną rolą demona jest rozgłoszenie informacji o dostępności tras. Niezbędne jest przesyłanie informacji o przynajmniej dwóch prefiksach. Pierwszy z nich dotyczy klasy przydzielonej do sieci lokalnej. Istotna jest skasowanie bitu *Autonomous*. Zgodnie ze specyfikacją konfiguracji bezstanowej [4], hosty nie będą mogły przydzielić sobie adresu z tej puli w trybie bezstanowym. Ważne jest również ustawienie bitu *OnLink*, który poinformuje hosty, że adresy z tej klasy są dostępne bezpośrednio i do ich osiągnięcia nie jest potrzebne wykorzystanie routera. Drugą rozgłaszaną trasą powinna być 2000::/3 ze skasowanymi flagami *Autonomous* oraz *OnLink*. Trasa ta jest trasą domyślną w rozwiązaniach IPv6. Dzięki tej informacji, hosty ruch do nieznanych sieci będą kierowały za pośrednictwem konfigurowanego właśnie routera.

3.3. Rozwiązywanie nazw

Dla końcowego użytkownika do pracy niezbędna jest usługa rozwiązywania nazw, czyli zamiana nazwy symbolicznej (np. *ksirouter.eti.pg.gda.pl*) na odpowiadający jej adres (np. *3ffe:8320:3:201::1*). Zadanie to może zostać zrealizowane poprzez klasyczną usługę DNS zrealizowaną w oparciu o protokół IPv6. Do tego celu bardzo dobrze nadaje się jedna z najbardziej popularnych otwartych implementacji protokołu DNS o nazwie *BIND*. Do prezentowanego tu rozwiązania konieczne jest użycie wersji 9. Obsługę nowego protokołu uzyskujemy za pomocą umieszczenia w pliku konfiguracyjnym wpisu:

```
listen-on-v6 {  
    any;  
};
```

Zamiast słowa kluczowego *any*, możliwe jest sprecyzowanie, na którym adresie ma być realizowana usługa.

3.4. Serwer DHCPv6

Kolejnym krokiem na drodze do wdrożenia IPv6 jest uruchomienie serwera protokołu DHCPv6. Konfiguracja zostanie omówiona na przykładzie otwartej implementacji o nazwie Dnsmasq [5].

Każdemu występującemu o konfigurację hostowi, niezbędne jest przekazanie przynajmniej dwóch informacji:

- adresu z obsługiwanej klasy
- adresu serwera rozwiązywania nazw
- domeny, w której znajduje się host (opcjonalnie)

Dodatkowo wskazane jest ustawienie odświeżania adresów co pewien okres. Zbyt mały okres spowoduje generowanie nadmiarowego ruchu związanego z przedłużaniem dzierżawy adresów, natomiast zbyt długi okres jest niewskazany w przypadku ewentualnej zmiany puli adresowej. Rozsądne wydaje się ustawienie minimalnego czasu odnawiania

adresów na 10 minut, natomiast całkowity czas życia adresu na 2 godziny. Podstawowa konfiguracja, spełniająca postawione tu założenia, wygląda następująco:

```
iface 'eth1'
{
  dns-server 2001:470:1f01::1
  domain eti.pg.gda.pl
  class
  {
    pool 2001:470:1f01::2-2001:470:1f01::ffff
    preferred-lifetime 3600
    valid-lifetime 7200
    T1 600
    T2 1200
  }
}
```

4. KONFIGURACJA HOSTÓW

Ostatnim krokiem na drodze do pełnego wykorzystania IPv6, jest konfiguracja hostów znajdujących się w sieci lokalnej. W porównaniu do poprzednich kroków, ta faza będzie powtarzana wielokrotnie. Typowa sieć wymaga konfiguracji jednego tunelu, jednego routera oraz setek hostów. Dlatego niezwykle istotne jest, aby ta faza była możliwie jak najkrótsza.

4.1. Uzyskanie adresu

Do pracy w sieci niezbędne jest uzyskanie adresu o zasięgu globalnym. Adres taki zostanie uzyskany za pomocą protokołu DHCPv6. W poprzednim rozdziale została opisana konfiguracja serwera DHCPv6. Na każdym z hostów musi zostać zainstalowany klient. Jednym z założeń implementacyjnych było tzw. podejście „out of the box”, czyli możliwość użytkowania od razu po instalacji, bez jakiegokolwiek dodatkowej konfiguracji. Dzięki temu czas potrzebny na konfigurację każdego hosta został ograniczony do minimum. Klient po uruchomieniu bez żadnej konfiguracji, odnajdzie serwer za pomocą komunikatu *SOLICIT*, a następnie wyśle prośbę o przydzielenie jednego adresu.

4.2. Konfiguracja routingu

Host musi wiedzieć, którędy ma wysyłać informacje adresowane poza sieć lokalną. W zależności od implementacji stosu IPv6 w systemie operacyjnym, przy starcie hosty mogą, ale nie muszą, wysyłać prośbę o „ujawnienie się” routerów (wiadomość ICMPv6, typ 133, tzw. *Router Solicit*). Router, po otrzymaniu takiej wiadomości, wysyła wiadomość *Router Advertisement* z informacją o dostępnych trasach. Jeżeli host w trakcie startu systemu nie wysyła *Router Solicit*, to konfiguracja routingu nastąpi w momencie okresowego rozsyłania wiadomości *Router Advertisement*, którą router co pewien ustalony czas generuje.

4.3. Rozwiązywanie nazw

Do korzystania z zasobów sieci, niezbędna jest poprawna konfiguracja usługi rozwiązywania nazw. Informacje o dostępnym serwerze DNS obsługującym protokół IPv6 zostaną dostarczone za pomocą opcji *DNS_Resolvers* zawartej we wiadomości *Reply*. Należy zwrócić uwagę na fakt, że konfiguracja tego parametru, podobnie jak wszystkich pozostałych po stronie hosta, następuje automatycznie. Eliminowana jest więc możliwość popełnienia błędu. Dodatkowo, użytkownik nie musi nawet zdawać sobie sprawy ze złożoności przeprowadzanych operacji.

4.4. Możliwości dodatkowe

Przedstawione rozwiązanie w żadnym wypadku nie wyczerpuje możliwości automatycznej konfiguracji w protokole IPv6. Przykładem dodatkowej opcji jest możliwość przekazania hostom informacji o tym, w jakiej domenie się znajdują. Znacznie ułatwia to korzystanie z zasobów sieci lokalnej. Przykładowo host znajdujący się w domenie *eti.pg.gda.pl* do połączenia się ze stroną wydziałową zamiast pełnej nazwy (*http:www.eti.pg.gda.pl*) może posłużyć się skróconą nazwą (*http:www*). W przypadku częstych połączeń między komputerami (np. laboratorium) znaczenie praktyczne takiego ułatwienia jest ogromne.

Do przekazania informacji o tym, w której domenie znajduje się host, służy opcja *Domain_list*. Jej użycie zostało opisane w punkcie przedstawiającym konfigurację serwera DHCPv6.

5. PRZEGLĄD SYSTEMÓW OPERACYJNYCH

Ogromna większość współczesnych systemów operacyjnych posiada zaimplementowany stos rodziny protokołów IPv6. Do wykorzystania przedstawionego tu rozwiązania konieczne jest wsparcie dla pewnych dodatkowych funkcji – DHCPv6 oraz obsługa wiadomości *Router Advertisement*. Oto wynik badań dla popularnych systemów operacyjnych.

Windows – Testowe wsparcie dla IPv6 zostało udostępnione dla platformy Windows 2000 [6]. Oficjalnie protokół ten jest wspierany w systemach XP i nowszych. Wiadomości *Router Advertisement* są poprawnie obsługiwane. Istnieje implementacja zarówno klienta, jak i serwera na systemy Windows. Twórcą jest Marek Senderski, absolwent wydziału ETI Politechniki Gdańskiej.

Linux – Wszystkie dostępne obecnie serie kerneli (2.4.x oraz 2.6.x) wspierają IPv6. Poprawna jest również obsługa wiadomości *Router Advertisement*. Istnieją 2 implementacje klienta i serwera protokołu DHCPv6. Twórcą jednej z nich jest Tomasz Mrugański, absolwent wydziału ETI Politechniki Gdańskiej i autor niniejszego opracowania.

BSD – IPv6 jest wspierane we wszystkich trzech odmianach BSD – FreeBSD, NetBSD oraz OpenBSD. Zarówno klient, jak i serwer są rozwijane a ramach projektu Kame. Według autorów, implementacja DHCPv6 powinna działać na wszystkich trzech odmianach systemu BSD.

Jak widać z powyższego zestawienia, wszystkie najczęściej używane systemy operacyjne, wspierają IPv6 i są gotowe do wykorzystania w proponowanym tu rozwiązaniu.

6. PODSUMOWANIE

Przedstawiona metoda szybkiego wdrażania IPv6 w sieciach lokalnych umożliwia szybkie przyłączenie sieci lokalnej do infrastruktury IP następnej generacji.

Wdrażanie IPv6 już teraz jest niezwykle istotne z psychologicznego punktu widzenia. W tej chwili obserwowany jest klasyczny problem jajka i kury. Użytkownicy końcowi nie domagają się realizacji nowych usług w technologii IPv6, ponieważ użycie IPv6 nie jest rozpowszechnione. Natomiast Producenci oprogramowania i usługodawcy rzadko oferują wsparcie dla IPv6, ponieważ popyt na tego typu usługi jest obecnie niewielki.

Należy zwrócić uwagę na fakt, że sednem prezentowanego tu rozwiązania jest wykorzystanie mechanizmu oferowanego przez protokół DHCPv6. Finalna wersja jego specyfikacji powstała w lipcu 2003 roku, jest to więc rozwiązanie stosunkowo nowe. Cieszy się ono dużym zainteresowaniem, gdyż od czasu publikacji standardu, wciąż pojawiają się propozycje nowych opcji rozszerzających. Przykładem takich opcji są *DNS Resolvers* oraz *Domain List*, które nie wchodzą w skład podstawowego protokołu.

Protokół DHCPv6 od samego początku był projektowany z myślą o dodatkowych opcjach. Dzięki modularnej budowie jego rozszerzanie jest trywialnie proste, a jego zakres zastosowania zdaje się nieograniczony.

BIBLIOGRAFIA

- [1] S.Deering, R.Hinden *Internet Protocol, Version 6 (IPv6) Specification*, RFC 2460, grudzień 1998
- [2] R.Droms, J.Bound, B.Volz, T.Lemon, C.Perkins, M.Carney *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*, RFC 3315, lipiec 2003
- [3] T.Narten, E.Kordmark, W.Simpson *Neighbor Discovery for IPv6*, RFC 2461, grudzień 1998
- [4] S.Thomson, T.Narten *IPv6 Stateless Address Autoconfiguration*, RFC 2462, grudzień 1998
- [5] T.Mrugalski, M.Senderski, *Dibbler – a portale DHCPv6*, <http://klub.com.pl/dhcpv6/>, luty 2004
- [6] K.Nowicki, J.Światowiak *Protokoły IPv6*, wyd. Politechniki Gdańskiej, 2001

RAPID IPv6 DEPLOYMENT METHOD USING DHCPv6 IMPLEMENTATION

Summary

Most important elements of next generation IP protocols has been published in the year 1995. Despite large capabilities, solutions based on IPv6 seem to appear fewer than expected. In Local Area Networks main issues are lack of solid service provider structure and complex configuration process. This essay analyses every phase required to gain IPv6 capability in LAN - tunneling methods, router and host configuration issues. Also rapid IPv6 deployment method is proposed. This method is based on DHCPv6 and various aspects of ICMPv6 protocols.

Jacek Rak

Katedra Systemów Informacyjnych, Politechnika Gdańska

PRZEŻYwalność Połączeń w Funkcji Częstości Awarii Elementów Sieci i Ich Położenia Geograficznego

Streszczenie

Artykuł przedstawia koncepcję zapewniania przeżywalności połączeń w sieciach szerokopasmowych. Pod pojęciem przeżywalności połączenia rozumiemy zdolność do przywrócenia realizacji połączenia (po wystąpieniu awarii elementu sieci) w czasie nie dłuższym od zadanego z góry, poprzez wyznaczenie lub uaktywnienie uprzednio wyznaczonej trasy alternatywnej. Rozpatrywana jest ochrona połączeń przed awarią pojedynczego łącza. Celem niniejszej pracy jest prezentacja autorskiego modelu zapewniania przeżywalności połączeń, który pozwala na jednoczesną redukcję wartości czasów odtwarzania połączeń, jak i stopnia wykorzystania łączy sieci. Rezultat ten jest bowiem niemożliwy do uzyskania przy zastosowaniu powszechnie stosowanego jednolitego dla całej sieci modelu zapewniania przeżywalności, gdyż wielkości te są ze sobą ujemnie skorelowane. Prezentowane podejście jest dedykowane przewodowym technologiom WDM i MPLS. Są one bowiem równoważne, jeśli utożsamimy etykietę MLPS z długością fali WDM.

1. WSTĘP

W sieciach szkieletowych zorientowanych połączeniowo, przed rozpoczęciem właściwego etapu transmisji danych, niezbędne jest zestawienie połączenia od krańca do krańca. Proces ten wymaga dokonania rezerwacji zasobów wzdłuż wyznaczonej ścieżki transmisji.

Ważnym elementem wpływającym na wiarygodność obsługi jest zdolność do szybkiego odtworzenia połączenia po awarii elementu sieci, który jest wykorzystywany przez ścieżkę transmisji. Badania pokazują, iż przeważającą część ogółu awarii stanowią uszkodzenia łączy podyktowane czynnikami zewnętrznymi.

Zdolność przywrócenia obsługi połączenia w czasie nie większym od z góry zadanego, określana mianem *przeżywalności*, jest możliwa dzięki wykorzystaniu dodatkowych tras – *ścieżek zabezpieczających*, będących alternatywnymi drogami dla podstawowej *ścieżki aktywnej* połączenia.

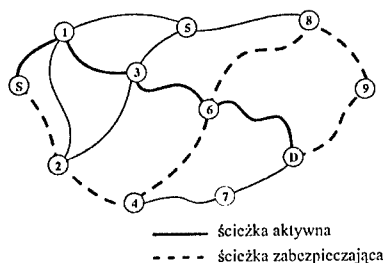
Istnieją dwa podstawowe podejścia wyznaczania ścieżek zabezpieczających, prezentowane w pracy [1]. W przypadku stosowania pierwszego z nich, określanego mianem *modelu reaktywnego*, wyznaczenie ścieżek zabezpieczających następuje dopiero po zaistnieniu awarii elementu sieci. Model ten nie daje jednak 100% gwarancji zapewnienia

przeżywalności połączenia, gdyż próba znalezienia ścieżki zabezpieczającej po wystąpieniu awarii może się skończyć niepowodzeniem z powodu niewystarczającego poziomu wolnych zasobów (np. przepustowości łączy sieci). Wartość czasu odtwarzania połączenia jest w tym przypadku również dodatkowo niekorzystnie powiększana o okres poszukiwania i instalacji ścieżki zabezpieczającej.

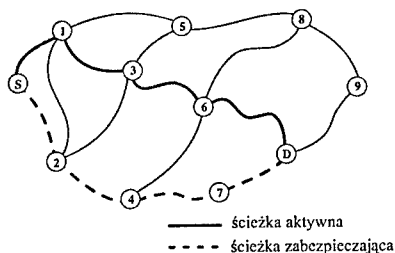
Podejściem wykorzystywanym w niniejszej pracy jest *uprzednio zaplanowane* wyznaczanie ścieżek zabezpieczających. Zakłada ono wyznaczanie i instalację zarówno ścieżki aktywnej, jak i ścieżek zabezpieczających danego połączenia przed rozpoczęciem właściwego etapu transmisji danych. Niemożność wyznaczenia choćby jednej z powyższych ścieżek, podyktowana niewystarczającym poziomem wolnych zasobów sieci, skutkuje odrzuceniem żądania zestawienia połączenia.

Konkretna ścieżka zabezpieczająca może chronić całą ścieżkę aktywną połączenia (model odtwarzania ścieżki – *ang. path protection*). Może również ochraniać jedynie pewien jej fragment. W tym ostatnim przypadku, dla danego połączenia niezbędne jest wyznaczenie kilku ścieżek zabezpieczających (omówione szerzej w rozdziale 1.1).

Często stosowane jest obecnie podejście wyznaczania tras zabezpieczających ścieżkę aktywną połączenia przed awarią pojedynczego węzła czy pojedynczego łącza. Zapewnienie ochrony ścieżki aktywnej przed awarią pojedynczego łącza wymaga, aby ścieżka zabezpieczająca była *rozłączna krawędziowo* z ochranianym przez nią obszarem ścieżki aktywnej. Z kolei, aby zapewnić ochronę przed awarią węzła, niezbędna jest *rozłączność węzłowa* ścieżki zabezpieczającej z ochranianym przez nią obszarem ścieżki aktywnej. Rozłączność węzłowa stanowi warunek silniejszy od rozłączności krawędziowej, gdyż awaria węzła sieci jest równoważna uszkodzeniu wszystkich łączy z nim incydentnych. Z tego powodu ścieżki zabezpieczające zapewniające przeżywalność połączenia na wypadek awarii węzła chronią również przed awarią łącza, ale nie na odwrót. Zarówno odmiany rozłączności ścieżek, jak i algorytmy ich wyznaczania zostały szerzej opisane w pracy [5].



Rys. 1. Połączenie (s,d) zrealizowane poprzez dwie rozłączne **krawędziowo** ścieżki



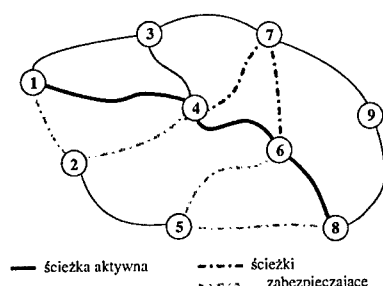
Rys. 2. Połączenie (s,d) zrealizowane poprzez dwie rozłączne **węzłowo** ścieżki

1.1. Zasięgi ochrony

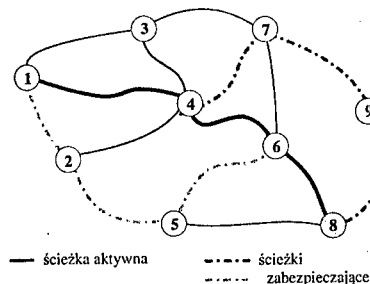
Biorąc pod uwagę wielkość obszaru ścieżki aktywnej połączenia ochranianego poprzez konkretną ścieżkę zabezpieczającą, powszechnie stosuje się następujące modele, prezentowane w [2, 3]:

- ochrony całej ścieżki aktywnej przez jedną ścieżkę zabezpieczającą (*ang. path restoration*). W przypadku modelu ochrony całej ścieżki, dla każdego połączenia wyznaczana jest tylko jedna ścieżka zabezpieczająca całą ścieżkę aktywną.
- ochrony pojedynczego łącza ścieżki aktywnej przez jedną ścieżkę zabezpieczającą (*ang. link restoration*) – rys. 3. W tym modelu każde łącze ścieżki aktywnej połączenia posiada własną ścieżkę zabezpieczającą.

Rozwiązaniem pośrednim pomiędzy dwoma pierwszymi podejściami jest autorska koncepcja ochrony, przez jedną ścieżkę zabezpieczającą, pewnego obszaru ścieżki aktywnej (*ang. region restoration*), przedstawiona na rys. 4. Rozmiar obszaru ochrony dla jednej ścieżki zabezpieczającej jest tutaj większy, niż w przypadku ochrony pojedynczego łącza ścieżki aktywnej (*link protection*), ale mniejszy niż dla modelu *path protection*.



Rys. 3 Przykład ścieżek zabezpieczających pojedyncze łącza ścieżki aktywnej



Rys. 4 Przykład ścieżek zabezpieczających obszary ścieżki aktywnej

W dalszej części pracy pojęcie „stopień wykorzystania zasobów” jest równoważne określeniu „stopień wykorzystania przepustowości łączy”. Im większe obszary ścieżek aktywnych są ochraniające przez ścieżki zabezpieczające, tym mniej potrzeba zasobów w celu zapewnienia przeżywalności połączeń. Najmniej zasobów potrzeba dla zapewnienia przeżywalności w przypadku stosowania modelu *path protection* – gdy ścieżka aktywna połączenia jest chroniona przez wyłącznie jedną ścieżkę zabezpieczającą. Odwrotna tendencja ma niestety miejsce biorąc pod uwagę wartości czasów odtwarzania. Im bowiem większy obszar ścieżki aktywnej chroni jedna ścieżka zabezpieczająca, tym wartości czasów odtwarzania są większe. Jest to spowodowane długościami ścieżek zabezpieczających oraz wykorzystywanymi protokołami odtwarzania połączenia, prezentowanymi w pracy [3].

Przy zastosowaniu dowolnego z zaprezentowanych powyżej modeli zapewniania przeżywalności, jednolitego dla wszystkich połączeń w sieci, niemożliwe jest jednoczesne zapewnienie niskich wartości czasów odtwarzania połączeń i małego stopnia wykorzystania zasobów. Pewnym rozwiązaniem może być tutaj koncepcja ochrony obszarowej, zapewniająca jednocześnie średnie wartości zarówno czasów odtwarzania połączeń, jak i stopnia wykorzystania zasobów sieci. Zależności wartości czasów odtwarzania połączeń i wykorzystania zasobów dla typowych wielkości obszarów ochrony zostały ukazane w tablicy 1.

Tablica 1

	odtworzenie ścieżki	odtworzenie obszaru	odtworzenie łącza
czas odtwarzania połączenia	duży	średni	mały
wykorzystanie zasobów	małe	średnie	duże

1.2. Metody optymalizacji

Zapewnienie przeżywalności połączeń, powodujące konieczność wyznaczania dodatkowych ścieżek zabezpieczających, prowadzi do zwiększenia stopnia wykorzystania zasobów sieci. Jest to spowodowane koniecznością rezerwacji przepustowości łącza na potrzeby ścieżek zabezpieczających. W celu ograniczenia tego zjawiska często stosuje się techniki optymalizacji polegające na zastosowaniu mechanizmów współdzielenia zasobów ścieżek zabezpieczających. Warunkiem niezbędnym do zastosowania współdzielenia zasobów zarezerwowanych dla ścieżek zabezpieczających jest, aby *wszystkie ochraniane przez te ścieżki fragmenty ścieżek aktywnych były ze sobą rozłączne* (odpowiednio rozłączne krawędziowo lub węzłowo). Rozłączność daje gwarancję, iż w wyniku wystąpienia awarii zaistnieje konieczność uaktywnienia wyłącznie jednej ścieżki z grupy ścieżek zabezpieczających, których zasoby podlegają współdzieleniu na danym łączu (bądź kilku łączach). Najpopularniejszymi odmianami optymalizacji stopnia wykorzystania zasobów ścieżek zabezpieczających są:

- *wewnętrzzadaniowa* – dotyczy możliwości uwspólnienia kanałów łącza ścieżek zabezpieczających rozłączne fragmenty ścieżki aktywnej tego samego połączenia
- *międzyzadaniowa* – określa zasady uwspólnienia kanałów łącza ścieżek zabezpieczających rozłączne fragmenty ścieżek aktywnych różnych połączeń
- *jednoczesna między- i wewnętrzzadaniowa* – stanowi połączenie obu powyższych technik. Jest ona zarazem najsilniejszą odmianą optymalizacji.

Każda z powyższych odmian optymalizacji, przeprowadzana w momencie wyznaczania wartości macierzy kosztów użycia łącza do budowy ścieżek (np. dla heurystycznego algorytmu Dijkstry) powoduje wzrost długości ścieżek zabezpieczających. Ścieżki te stają się tym dłuższe, im silniejsza jest optymalizacja. Wyznaczając ścieżki zabezpieczające, algorytm heurystyczny wybiera bowiem znacznie częściej nie łącza, które są najkrótsze w sensie stosowanej metryki, ale te, dla których można przeprowadzić optymalizację. Stosowanie optymalizacji prowadzi więc do powiększenia różnicy pomiędzy wartościami czasu odtwarzania połączeń a stopniem wykorzystania zasobów.

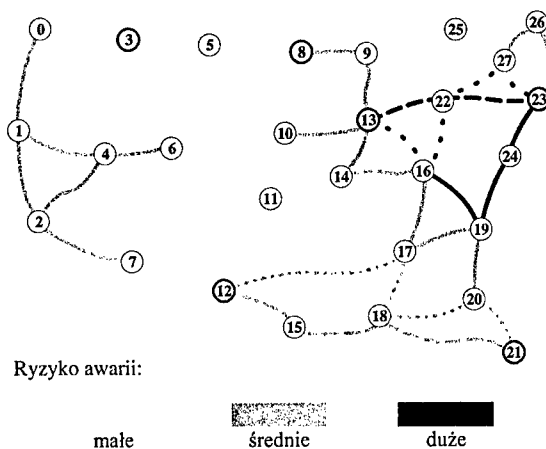
2. ZAŁOŻENIA KONCEPCYJNE MODELU

Optymalizacja stopnia wykorzystania zasobów, polegająca na dokonywaniu współdzielenia kanałów łącza ścieżek zabezpieczających (opisana w rozdziale 1.2), pomimo iż powoduje redukcję stopnia wykorzystania zasobów, prowadzi do zwiększania czasów odtwarzania połączeń. Z tego powodu koncepcja ta nie jest w omawianym modelu brana pod uwagę. Celem prezentowanego podejścia jest uzyskanie bowiem jak najmniejszych wartości zarówno średniego stopnia wykorzystania zasobów sieci, jak i średniego czasu odtwarzania połączeń.

Ideą koncepcji prezentowanej w niniejszej pracy jest uzależnienie w sposób dynamiczny wielkości obszarów ochrony ścieżki aktywnej połączenia od wartości współczynników awaryjności elementów sieci, wynikających z położenia geograficznego. Współczynniki te określają poziom ryzyka wystąpienia awarii określonych elementów sieci (tutaj łączy).

- Dla łączy o małym lub w ogóle pomijalnym ryzyku awarii elementu sieci, podyktowanym czynnikiem zewnętrznym, obszary odtwarzania mogą być duże. Pojedyncza ścieżka zabezpieczająca może chronić w skrajnym przypadku nawet całą ścieżkę aktywną. Prowadzi to do uzyskania niskiego stopnia wykorzystania zasobów, bowiem odtwarzanie o zasięgu całej ścieżki aktywnej jest najbardziej oszczędną formą zapewniania przeżywalności połączeń. W proponowanym modelu omawiane podejście jest stosowane dla łączy o najniższej wartości współczynnika awaryjności łączy. Ewentualna awaria zdarza się w tych obszarach rzadko, toteż uzyskiwane dla nich duże wartości czasu odtwarzania połączeń nie wpływają w sposób znaczny na jego wartość średnią dla całej sieci.

- Im większe ryzyko awarii, tym poszczególne obszary ochrony ścieżki aktywnej należy zmniejszać, aż do uzyskania w skrajnym przypadku sytuacji ochrony pojedynczego łącza ścieżki aktywnej przez jedną ścieżkę zabezpieczającą. Krótki czas odtwarzania oraz relatywnie większa częstość występowania awarii w obszarach stosowania krótkich ścieżek zabezpieczających sprawiają, iż osiągane małe wartości czasów odtwarzania połączeń mają duży wkład w jego wartość średnią dla całej sieci, która tym sposobem jest niska. Im mniejszy obszar ochrony, tym stopień wykorzystania zasobów wzrasta, lecz jeśli obszarów ochrony o dużym ryzyku nie będzie w sieci dużo, to stopień ten będzie pozostawał na akceptowalnym poziomie.



Rys. 5. Trzystopniowa skala awaryjności łączy używana do określenia wielkości obszarów ochrony

Prezentowana koncepcja została zilustrowana na rys. 5. Przedstawione są trzy połączenia, odpowiednio pomiędzy węzłami 3 i 8, 13 i 23 oraz 12 i 21. Na rys. 5 wartości współczynników awarii przekładają się na trzystopniową skalę ocen ryzyka: (małe, średnie i wysokie). O przynależności do odpowiedniej klasy świadczy kolor łącza (jasnosiwy, ciemносиwy, czarny).

Ścieżka aktywna (3-5-8) połączenia (3,8) wykorzystuje łącza o niewielkim ryzyku awarii. Można więc zastosować dla tego połączenia model *path restoration*, czyli wyznaczyć jedną ścieżkę zabezpieczającą całą ścieżkę aktywną. Uzyskuje się w ten sposób najmniejszy (w porównaniu z modelami odtwarzania obszarowego i odtwarzania łącza) stopień wykorzystania zasobów w celu zapewnienia przeżywalności danego połączenia.

Awaria elementów sieci, które są wykorzystywane przez ścieżkę aktywną połączenia (3,8) jest mało prawdopodobna, toteż duża wartość czasu odtwarzania po ewentualnej awarii ma mały wpływ na jego wartość średnią dla całej sieci.

Ścieżka aktywna (13-22-23) połączenia (13,23) z rys 5 wykorzystuje łącza o najwyższej wartości współczynnika awaryjności, toteż niezbędne jest w tym przypadku zastosowanie możliwie jak najkrótszych ścieżek zabezpieczających. Zastosowanie ochrony każdego łącza ścieżki aktywnej przez oddzielną ścieżkę zabezpieczającą wydatnie skraca wartość średniego czasu odtwarzania połączeń. Jest to szczególnie ważne, gdyż awarie występują w tym przypadku relatywnie często.

Dla ostatniego przykładowego połączenia (12,21), ścieżka aktywna (12-15-18-21) wykorzystuje łącza o średnich wartościach współczynnika awaryjności. Odpowiednie wydaje się więc zastosowanie rozwiązania pośredniego, czyli odtwarzania obszarowego, (tutaj dzięki ścieżkom (12-17-18) i (18-20-21), zabezpieczającym odpowiednio obszary (12-15-18) i (18-21) ścieżki aktywnej).

Zarówno ścieżki aktywne, jak i zabezpieczające są wyznaczane przy użyciu heurystycznego algorytmu Dijkstry. W celu określenia kosztów użycia łączy wykorzystywana jest metryka długości łącza.

Ścieżka aktywna i ścieżki zabezpieczające każdego połączenia są wyznaczane przed rozpoczęciem właściwego etapu transmisji danych. Wystąpienie awarii elementu, który jest wykorzystywany przez ścieżkę aktywną, implikuje konieczność uaktywnienia ścieżki zabezpieczającej i wymusza konieczność wyznaczenia nowego zestawu ścieżek: aktywnej i zabezpieczających danego połączenia. W stanie awarii (czyli od momentu wykrycia uszkodzenia do chwili wyznaczenia nowych ścieżek obsługi), w celu obsługi ruchu wykorzystywana jest wyznaczona uprzednio (przed wystąpieniem awarii) ścieżka zabezpieczająca.

Co pewien przedział czasu niezbędna jest aktualizacja ścieżek zabezpieczających. Dotyczy ona również tych połączeń, których ścieżki aktywne nie uległy awarii. Jest to związane z zmieniającymi się w czasie wartościami współczynników awaryjności elementów sieci i wynikającą z tego koniecznością dostosowania postaci ścieżek zabezpieczających do bieżących wartości współczynników. Opisane postępowanie ma za zadanie utrzymać optymalną wartość stosunku czasów odtwarzania połączeń i stopnia wykorzystania zasobów sieci w obliczu zmieniających się w czasie współczynników awaryjności elementów sieci.

3. OPIS PRZEPROWADZONYCH BADAŃ

Badania zostały przeprowadzone przy użyciu napisanego w tym celu symulatora. Siecią podlegającą badaniu była amerykańska szerokopasmowa *U.S. Long-Distance Network* (prezentowana w pracy [6], zawierająca 28 węzłów i 45 łączy) – rys. 5. Przeprowadzono pomiary wartości czasów odtwarzania połączeń i stopnia wykorzystania zasobów dla dwóch scenariuszy wartości współczynników awaryjności elementów sieci. Oba scenariusze zakładały istnienie trzech klas awaryjności (ryzyko małe, średnie i duże). Scenariusz A zakładał początkowy przydział poszczególnych łączy sieci do klas awaryjności zgodnie z rys. 5. Natomiast w przypadku scenariusza B, łącza należały początkowo do dwóch klas – jedynie o ryzyku najmniejszym i największym. Te łącza, które w scenariuszu A cechowało początkowo średnie ryzyko awaryjności, w przypadku scenariusza B miały ryzyko małe. Klasy awaryjności pozostałych łączy w scenariuszu B były natomiast takie same jak w przypadku scenariusza A.

Każdy pomiar polegał na wygenerowaniu 30 topologii logicznych. Każda z topologii logicznych była grafem żądań zestawienia 30 losowych połączeń. W przypadku pomiaru wartości czasów odtwarzania połączeń T_R , dla każdej topologii logicznej generowano awarie 30 losowo wybranych łączy. Celem było zbadanie iloczynu wartości średniej czasu odtwarzania połączeń (μ_{TR}) i wartości średniej stopnia wykorzystania przepustowości łączy (μ_U) dla każdego modelu (tablica 2). Im mniejsza wartość iloczynu $\mu_{TR} \cdot \mu_U$, tym osiągnięte rezultaty są lepsze. Badaniom poddany został model ochrony połączeń przed awarią pojedynczego łącza.

4. ANALIZA REZULTATÓW

SCENARIUSZ A

Biorąc pod uwagę wykorzystanie procentowe zasobów, proponowany model jest porównywalny z koncepcją odtwarzania obszarowego. Wartość średnia czasu odtwarzania połączeń jest natomiast zbliżona do najlepszych wyników osiąganych w przypadku koncepcji odtwarzania łącza (z powodu dużej częstości występowania awarii elementów sieci w rejonach, w których stosowane jest właśnie odtwarzanie łącza). Iloczyn rozpatrywanych czynników jest więc znacznie mniejszy niż w przypadku jakiegokolwiek jednolitego dla całej sieci modelu odtwarzania.

SCENARIUSZ B

Dla scenariusza zakładającego występowanie początkowo łączy o jedynie najwyższym i najniższym współczynniku awaryjności, osiągnięto lepsze rezultaty zarówno w dziedzinie czasów odtwarzania połączenia, jak i stopnia wykorzystania zasobów. Z tego powodu wartość iloczynu rozpatrywanych składników staje się jeszcze mniejsza, niż w scenariuszu A – tutaj 537,97.

Tablica 2

	Path restoration	Region restoration	Link restoration	PROPONOWANY MODEL	
				Scenariusz A	Scenariusz B
μ_{T_R} [ms]	47,92	43,79	23,98	26,55	26,09
μ_U [%]	18,30	20,01	30,56	20,74	20,62
$\mu_{T_R} \cdot \mu_U$ [ms*%]	876,93	876,24	732,83	550,65	537,97

5. PODSUMOWANIE

Wyniki otrzymane dla proponowanego modelu na podstawie badań symulacyjnych są zgodne z oczekiwaniami. Uzyskano znacznie mniejsze wartości iloczynu $\mu_{TR} \cdot \mu_U$ niż w przypadku stosowania modeli odtwarzania, jednolitych dla wszystkich połączeń w sieci. Proponowana koncepcja nadaje się szczególnie do zastosowania w przypadku, gdy współczynniki awaryjności łączy przyjmują wyłącznie wartości najniższe i najwyższe

(scenariusz B - tablica 2). W tym przypadku iloczyn wartości średnich czasu odtwarzania połączenia i stopnia wykorzystania zasobów osiąga poziom najniższy spośród wszystkich rozpatrywanych podejść.

BIBLIOGRAFIA

- [1] Kawamura R.: *Architectures for ATM network survivability*, IEEE Communications Surveys Fourth Quarter 1998 vol.1 no.1
- [2] Ramamurthy S., Mukherjee B.: *Survivable WDM Mesh Networks, Part I – Protection*, Proc. IEEE INFOCOM, 03.1999, s. 744-751
- [3] Ramamurthy S., Mukherjee B.: *Survivable WDM Mesh Networks, Part II – Restoration*, Proc. IEEE Integrated Circuits Conf., 06.1999, s. 2023-2030
- [4] Suurballe J.W.: *Disjoint Paths in a Network*, Networks, John Wiley & Sons 1974, s.125-145
- [5] Bhandari R.: *Survivable networks – algorithms for diverse routing*, Kluwer Academic Publishers, 1999
- [6] Xiong Y.: *Restoration strategies and spare capacity requirements in self-healing ATM Networks*, IEEE/ACM Transactions on Networking, 02. 1999, vol. 7, no. 1

SURVIVABILITY OF CONNECTIONS IN THE FUNCTION OF NETWORK ELEMENTS' FAILURE FREQUENCY AND THEIR GEOGRAPHIC LOCATION

Summary

This article presents an idea of assuring survivability of connections in broadband networks. Survivability means an ability to restore a connection (after a failure of a network element) within a certain maximum amount of time by finding or activating a previously found backup path. The model presented here is to assure the survivability of connections in case of a single link failure. The aim of this article is to present an innovative model of assuring survivability of connections which allows both to reduce the network resource usage and the values of restoration times. This result is impossible to be achieved when using a uniform model of assuring survivability of all the connections in the network. It's mainly because restoration time values and network resource usage are negatively correlated. Approach presented here is dedicated to WDM and MPLS technologies. These two technologies are equivalent if we identify an MPLS label with a WDM wavelength.

Lech Smoleński

Katedra Systemów i Sieci Telekomunikacyjnych, Politechnika Gdańska

**KIERUNKI ROZWOJU TECHNIK TRANSMISJI
W WIELOUSŁUGOWYCH OPTYCZNO-MIEDZIANYCH
SIECIACH DOSTĘPOWYCH**

Streszczenie

Referat przedstawia kierunki zmian w sieciach dostępowych, wykorzystujących linie miedziane i światłowodowe, prowadzące do uzyskania sieci oferujących pełny zestaw usług szeroko- i wąskopasmowych (FSAN). Celem takiej sieci jest udostępnianie jej użytkownikom usług, konwencjonalnych i nowych, z właściwym dla sieci telekomunikacyjnej poziomem jakości i niezawodności. Zmiany w sieciach dostępowych polegają na rosnącym udziale transmisji optycznej, rozpowszechnieniu transmisji o asymetrycznej przepływności dla linii miedzianych (ADSL), dążeniu do uzgodnienia technik transmisji i protokołów z sieciami szkieletowymi i metropolitalnymi (ATM, IP, warstwa optyczna). Kolejne warianty rozwojowe sieci dostępowej zmierzają do pełnej integracji, gdy wszystkie usługi będą realizowane przez wspólny węzeł usługowy.

1. WSTĘP

W sieci następnej generacji (NGN), niezależnie od stosowanych protokołów, będzie niezbędna ujednolicona transmisja informacji cyfrowej w warstwie fizycznej. Rozwój infrastruktury informacyjnej wymaga rozwiązania, którego celem ma być oferowanie użytkownikom sieci komunikacyjnej usług dostępnych we wcześniej istniejących sieciach, z właściwym dla sieci telekomunikacyjnej poziomem jakości i niezawodności, oraz zapewnienie współpracy różnych sieci transportowych. W wypadku sieci dostępowych zmiany polegają na rozpowszechnieniu transmisji asymetrycznej pod względem przepływności (systemy ADSL/VDSL dla linii miedzianych), rosnącym udziale transmisji pakietowej (ATM lub IP), dążeniu do uzgodnienia technik transmisji i protokołów z sieciami szkieletowymi i metropolitalnymi (IP, warstwa optyczna).

Wielousługowy charakter sieci dostępowej powoduje dążenie do realizacji wszystkich usług w sposób możliwie jednorodny, korzystając na poziomie transmisji fizycznej z tego samego systemu, którego przepustowość powinna umożliwiać równoległą realizację wielu usług, w tym również szerokopasmowych. Podobny efekt wzrostu wymaganej przepustowości powoduje coraz częstsze korzystanie z sieci dostępowej przez niewielkie sieci lokalne (domowe lub SOHO), w których równolegle korzysta z usługi dostępu do Internetu kilku użytkowników.

Na obszarze obejmowanym przez sieć dostępową współistnieją obecnie jako medium transmisyjne dla przesyłanych sygnałów włókna światłowodowe i linie miedziane (symetryczne w sieci wykorzystywane przez PSTN, współosiowe w sieciach CATV). Tendencja do przybliżania zakończeń optycznych do użytkownika występuje nadal, jednak na masową skalę wykorzystywane są rozwiązania hybrydowe FITL, a doprowadzenie sygnałów optycznych do każdego użytkownika POTS/ISDN (FTTH/O) jest na dziś nieekonomiczne. Intensywnie rozwijane techniki dostępu radiowego, w postaci WLAN mogą upowszechnić się jako włączone w schemat hybrydowej sieci dostępowej, zastępując ostatni odcinek linii miedzianej w bezpośrednim sąsiedztwie użytkowników, w połączeniu z dalszymi fragmentami sieci realizowanymi w oparciu o transmisję optyczną lub xDSL. Ze względu na ograniczenia pasma radiowego, które występują pomimo udostępnienia zakresu fal centymetrowych i postępów techniki odbioru radiowego, szerokopasmowy dostęp radiowy (np. LMDS) pokrywający pełny zasięg sieci dostępowej może być wykorzystywany przez stosunkowo niewielką grupę użytkowników.

Konwencjonalne sieci dostępowe, projektowane dla realizacji usług POTS/ISDN, oparte na transmisji w trybie łączy, osiągnęły swoją dojrzałość techniczną jako sieci FITL, wykorzystujące mechanizmy transmisji optycznej dla dużych wiązek kanałów, zwielokrotnianych w technice SDH (czasami dla mniejszej liczby kanałów także PDH). Na odcinku „ostatniej mili” w sąsiedztwie użytkowników powszechnie jest wykorzystywana transmisja w symetrycznych liniach miedzianych, zgrupowanych w kablach o konstrukcji czwórkowej, typowych dla dostępu analogowego POTS [1]. W takiej sieci koegzystują interfejsy analogowe dla POTS i cyfrowe dla ISDN, uzupełnione systemami xDSL w wypadku realizacji dostępu z większą szybkością niż ISDN-BRA.

2. ZASADY FUNKCJONOWANIA WIELOUSŁUGOWYCH SIECI DOSTĘPOWYCH

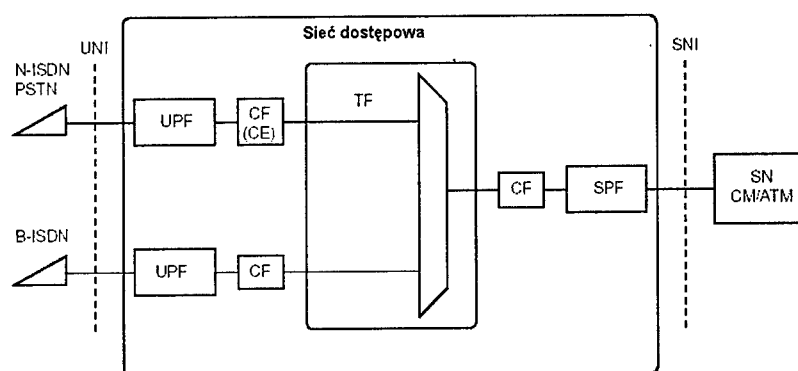
Dla sieci dostępowej typowa jest konfiguracja typu punkt-wiele punktów, która wymaga zwielokrotnienia przy przesyłaniu informacji w dół a przy transmisji w górę odpowiedniego mechanizmu wielodostępu. Niektóre nowe usługi wymagają przepływności większych niż udostępniane przez PSTN i ISDN, takie usługi są najlepiej realizowane w sieci całkowicie optycznej lub hybrydowej optyczno-miedzianej. Gdy usługi ewoluują i są wprowadzane nowe usługi, przepływności i wymogi dotyczące zarządzania wzrastają. Wymaga to od sieci dostępowej elastyczności i łatwości rozszerzania i modernizacji.

W ramach wielousługowej sieci dostępowej można wykorzystywać do przesyłania sygnałów w trybie łączy, równolegle z sygnałami w trybie pakietowym, wspólną warstwę transportową opartą na zwielokrotnieniu kanałów (np. SDH) lub na transmisji pakietowej. W tym ostatnim przypadku potrzebna jest emulacja kanału w sieci pakietowej (np. ATM). Interfejsy węzłów usługowych (SNI) mogą być rozdzielone lub SNI może być wspólny. Obecnie za docelową strukturę sieci dostępowej przyjmuje się schemat z jednym węzłem usługowym dla wszystkich usług, zarówno w trybie kanałów jak i pakietów. Na podstawie [2] można przedstawić warianty konfiguracji pozwalające na integrację dostępu PSTN ISDN i szerokopasmowego (B-ISDN) we wspólnej sieci dostępowej przy użyciu różnych warstw sieciowych.

W wariantcie typowym dla użytkowanych obecnie sieci dostępowych multipleksacja w warstwie mediów transmisyjnych, może być zastosowana w warstwie fizycznej (światłowod, para miedziana, tor współosiowy, transmisja radiowa) lub w warstwie sekcji (ścieżki

SDH lub PDH) w celu integracji dostępu w trybie łączy i ATM. Nie wpływa to na funkcjonalność sieci dostępowej a SNI i UNI dzielą to samo medium transmisyjne.

Multipleksacja w warstwie ścieżek z integracją trybu łączy i ATM, stosując zwielokrotnienie w warstwie ścieżek dla ATM, wymaga konwersji informacji przesyłanej w trybie łączy, przez użycie funkcji emulacji łączy przed połączeniem z informacją ATM. Informacja jest demultiplesowana przed doprowadzeniem do węzłów usługowych.



Rys.1. Pełna integracja w węźle usługowym dla usług POTS, N-ISDN i szerokopasmowych (B-ISDN) dla FSAN [2].

W kolejnym wariancie integracja dla trybu łączy i ATM, zakłada multipleksację w warstwie ścieżek (ATM), jednak zbiorczy strumień informacji nie jest demultiplesowany w sieci dostępowej a jest przekazywany przez wspólny SNI do SN. Demultipleksacja strumienia informacji jest przeprowadzana w węźle usługowym ATM przy użyciu funkcji emulacji łączy dla transferu ATM, przed wprowadzeniem do SN w trybie łączy.

W wypadku pełnej integracji (rys.1) węzeł usługowy ATM demultiplesuje zbiorczy strumień informacji i zarządza obu trybami transferu informacji wewnętrznie.

3. TRANSMISJA OPTYCZNA W SIECI PASYWNEJ PON

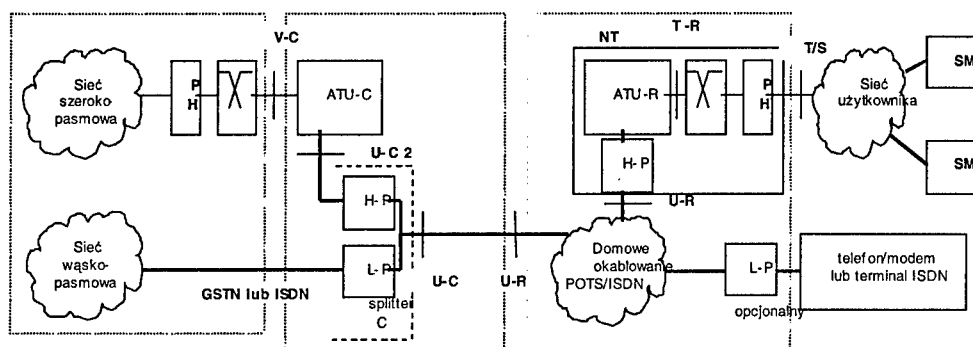
Pasywna sieć optyczna PON jest uważana obecnie za najlepsze rozwiązanie dla realizacji dostępu szerokopasmowego w sieci FITL, pozwalając na doprowadzenie sygnału optycznego do ONU, zlokalizowanego u użytkownika (wariant FTTH/O) lub podział zasobów ONU między większą liczbę użytkowników w wariantach FTTC/B. Jednym z efektów zastosowania sieci PON w obszarze dostępowym jest ujednolicenie techniki transportu optycznego z sieciami szkieletowymi. Sieć PON, niezależnie od konkretnych rozwiązań i wyboru standardów dla przesyłanych danych (A-PON, E-PON, G-PON) jest najkorzystniejszym rozwiązaniem dla zapewnienia pełnej gamy usług video (w tym prawdziwe VoD, HDTV).

Grupa zaleceń G.983.x definiuje rodzinę sieci ATM-PON, dostosowaną do usług szerokopasmowych, wykorzystującą transmisję ATM ale nie wykluczającą innych technik transmisji. Transmisja w wersji symetrycznej może następować z nominalną szybkością 155.52 Mb/s lub 622.08 Mb/s; w wariancie niesymetrycznym 622.08 w dół/155.52 w górę. Klasy dla ścieżek optycznych w sieci PON są zróżnicowane pod względem wartości minimalnych i maksymalnych tłumienności dla takich ścieżek (klasa A tylko w G.983.3). Dla

4. LINIE SYMETRYCZNE MIEDZIANE I ICH WYKORZYSTANIE DLA SYSTEMÓW DSL

Sieć miedziana, funkcjonująca na odcinku „ostatniej mili” i dostosowana do transmisji analogowej w pasmie akustycznym, jest już od kilkunastu lat używana do transmisji cyfrowej a kolejne wersje standardów korzystają z coraz bardziej zaawansowanych technologii DSP, pozwalając na widoczną poprawę parametrów. Wiąże się z tym opracowywanie kolejnych generacji standardów, charakteryzujących się coraz większą przepływnością maksymalną, wzrostem zasięgu transmisji dla danej przepływności a także nowymi cechami funkcjonalnymi, między innymi elastycznością konfiguracji i zdolnością do adaptacji do zmian warunków w linii w trakcie pracy.

Ewolucja usług poza znane powszechnie usługi sieciowe (email, FTP, www, chat) prowadzi do wzrostu wymagań dotyczących przepływności. Kompresja video, coraz skuteczniejsza w redukcji nadmiarowości, jest czynnikiem wpływającym na upowszechnienie systemów ADSL, oferujących przepływności pozwalają na przenoszenie kilku kanałów równolegle lub udostępnianie TV wysokiej jakości (HDTV). Ułatwia to konkurencję z sieciami TV kablowej, które mogą z łatwością zapewnić dostęp szerokopasmowy w sieci HFC, ale na ogół sprzęt wymagany do realizacji takiego dostępu jest droższy niż w wypadku ADSL.



Rys.3. Model referencyjny dla typowej konfiguracji systemu ADSL zgodnego z G.992.2 do 4 (bez splitterów po stronie użytkownika) [5].

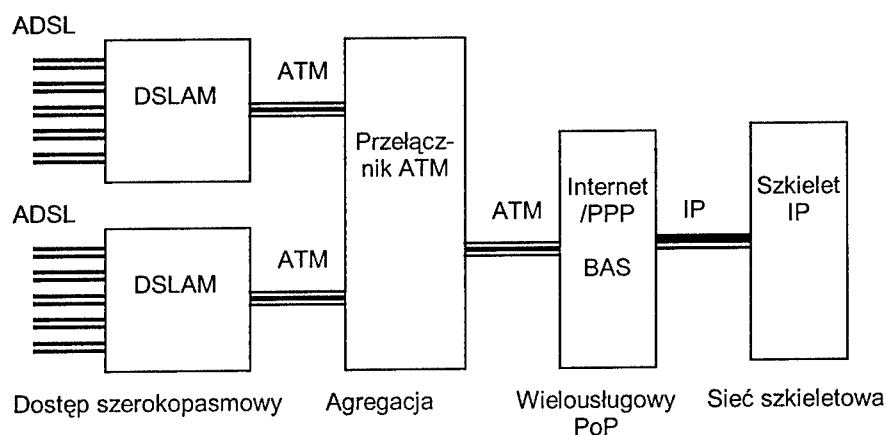
Kolejne wersje standardów dotyczących transmisji ADSL uwzględniają dążenie do wzrostu przepływności przypadającej na jedną linię a także do zapewnienia dostępu dla maksymalnej części użytkowników miedzianej sieci dostępowej. Nie jest możliwa maksymalizacja równocześnie obu tych wielkości, natomiast typowa jest możliwość wymiany zasięgu transmisji w linii miedzianej na przepływność.

Stopniowanie wymogów dotyczących przepływności następuje od wartości dla podstawowej usługi (0.5 do 1.5 Mb/s) ze skokiem co 2 Mb/s dla asymetrycznego strumienia („w dół” do użytkownika) w zależności od liczby urządzeń korzystających równocześnie z dostępu u pojedynczego użytkownika [6].

W typowych sieciach dostępowych opartych na parach miedzianych, użytkownicy są zlokalizowani w odległości od kilkuset metrów do ok. 2 km („ostatnia mila”) od szafki rozdzielczej, do której jest doprowadzony wieloparowy kabel magistralny lub w nowszych

realizacjach jako FITL, światłowodowy system transmisyjny zakończony ONU. Pojawia się tendencja do przystosowania każdego z nowo instalowanych zakończeń abonenckich dla POTS do równoległej transmisji ADSL. Wiąże się to z potrzebą modyfikacji fragmentów miedzianych sieci dostępowych, które są zbyt oddalone obecnie od węzłów oferujących dostęp ADSL, tak aby transmisja była możliwa dla najdłuższych linii w sieci.

Wzrost dostępności wymaga wprowadzenia wyposażenia ADSL w punktach wyniesionych, takich jak ONU w sieci FTTC. Wiąże się to z wprowadzeniem modułów DSLAM, dokonujących multipleksacji statystycznej, dla lepszego wykorzystania zasobów transmisyjnych sieci dostępowej. Wzrost przepływności, związany z poszerzeniem zakresu usług, może wymagać od dostawców usług wprowadzenia urządzeń ADSL w punktach wyniesionych w obszarze wcześniej obsługiwanym, ale z niewystarczającą szybkością. Oba te kierunki powodują upowszechnienie wyniesionych urządzeń ADSL w sieci FITL.



Rys.4. Agregacja sygnałów dostępu ADSL w sieci dostępowej przy użyciu przełącznika ATM [7].
Analogiczny schemat można przedstawić dla standardu Ethernet.

Aktualne standardy dla miedzianych sieci dostępowych przewidują wykorzystanie zarówno transmisji symetrycznej pod względem przepływności (HDSL: G.991.1; SHDSL: G.991.2; VDSL: G.993.1) jak i niesymetrycznej (ADSL I generacji: G.992.1 i wersja uproszczona G.992.2; ADSL2: G.992.3 i wersja uproszczona G.992.4; ADSL2+: G.992.5; VDSL: G.993.1). W obszarze sieci miedzianej, opartej na parach symetrycznych następuje przede wszystkim upowszechnienie modemów ADSL, wykorzystujących obecnie technikę ATM, dostosowanych do realizacji usług związanych z asymetryczną przepływnością do i od sieci. Uzupełnienie dla małych odległości i dużych szybkości stanowią systemy VDSL.

Pierwsza grupa standardów ADSL [5, 8] jest stosunkowo mało elastyczna i w wersji europejskiej zawiera trzy podstawowe warianty różniące się szybkością transmisji w dół w kanałach simpleksowych (6, 4 lub 2 Mb/s) przy małym, średnim lub maksymalnym zasięgu. Istotną cechą jest możliwość równoczesnej transmisji po tej samej parze analogowego sygnału telefonicznego w pasmie naturalnym a w kolejnej wersji sygnału dla dostępu ISDN-BRA (punkt U). Wymaga to zastosowania rozdziału częstotliwościowego pasma przy pomocy filtrów (splitterów) na końcach linii miedzianej - obu lub tylko jednym (w wersji lite G.992.2/4). Wersja uproszczona o mniejszych wymaganiach udostępnia też mniejsze przepływności w transmisji w obu kierunkach, ale jest prostsza i tańsza.

Tablica 3

Aktualny stan standaryzacji dla systemów ADSL

Zalecenie ITU-T	Nazwa komercyjna	Max. częst. pasma lin.	Max. prz. „w dół”	Max. prz. „w górę”	Zasięg ($\alpha=0.5$)	dla przepływności
		MHz	Mb/s	kb/s	km	Mb/s
G.992.1	ADSL	1.1	6.144	640	4.2	0.576
G.992.2	ADSL Lite	0.55	1.536	0.512	4.2	0.512
G.992.3	ADSL2	1.1	8.0	800	5.5	0.288
G.992.4	ADSL2 Lite	0.55	1.536	0.512	4.2	0.512
G.992.5	ADSL2+	2.2	16	1.5	5.5	0.33

Kolejna para standardów [8], znana jako ADSL2, uwzględnia aktualne możliwości techniczne, w szczególności DSP, i dzięki temu pozwala na uzyskanie większej szybkości transmisji przy danej długości linii lub większego zasięgu dla danej przepływności. Zalety ADSL2 to m.in. zredukowana objętość nagłówka, możliwość zmian położenia sygnału pilota z punktu widzenia odbiornika, negocjacja szybkości transmisji, obowiązkowe kodowanie kratowe, umożliwienie kasowania interferencji radiofonicznych, regulowany przedział czasowy inicjalizacji, dodatkowy stan transmisji z małą mocą, rozbudowana diagnostyka linii (testowanie z obu końców i przysyłanie wyników z odległego końca, szum tła i S/N w funkcji częstotliwości).

Istotnymi usprawnieniami są: płynna adaptacja przepływności do warunków w linii, dynamiczny podział przepływności, do 4 równoległych kanałów podkładowych, praca wieloparowa, nowe tryby całkowicie cyfrowe (32 lub 64 kanały „w górę”), symetryczna transmisja 1 Mb/s na odległość do 4 km, większe przepływności „w górę”, ustawialny kształt widma podczas pracy i inicjalizacji, rozdział częstotliwościowy dla POTS/ISDN bez splittera po stronie użytkownika.

Zal. G.992.3 definiuje różne kombinacje usług POTS/ISDN realizowanych w dolnym pasmie przy zastosowaniu rozdziału częstotliwościowego lub bez nich. Wiąże się to z rozdziałem pasma częstotliwości pomiędzy usługi POTS/ISDN i ADSL a równocześnie powoduje konieczność odpowiedniego ograniczenia pasma i gęstości widmowej dla transmisji ADSL. Modem może pracować w trybie cyfrowym bez dodatkowych usług a inną opcją jest zapewnienie wyodrębnionej przepływności kanału dla POTS lub ISDN w łączu DSL. Wariant przysyłania sygnałów POTS/ISDN w postaci cyfrowej przez DSL jest w tym wypadku korzystniejszy z obu wyżej wymienionych względów. Zgodność G.992.3 z innymi standardami ADSL zostaje zachowana (zarówno z wcześniejszymi jak i z wersją uproszczoną G.992.4) dla podstawowych funkcji, natomiast zaawansowane opcje wymagają zgodności co do wariantu standardu dla ATU-R i ATU-C. Podstawowe różnice dla wersji lite (G.992.4) to 128 nośnych dla DMT (jak G.992.2), zdolność do pracy powyżej pasma POTS (ale nie ISDN) i bez kanału analogowego przy 32 kanałach w górę.

Kolejny wariant standardu ADSL2+ (G.992.5 z 2003 r.) jest to wersja ADSL2, udostępniająca większą przepływność w dół, przede wszystkim dla linii o mniejszej długości, (przykładowo 16 Mb/s dla linii o długości 1300m, powyżej 2.5 km wartości nie odbiegają od uzyskiwanych dla ADSL2). Uzyskuje się to dzięki dwukrotnemu poszerzeniu pasma liniowego (do 2.2 MHz). Dla mniejszych szybkości (300/64 kb/s) możliwa jest transmisja na odległość do 6 km. Wprowadzono możliwość testowania linii z jednego końca przy braku komunikacji, zarówno dla ATU-R jak i ATU-C, dla potrzeb lokalizacji uszkodzeń (np. przerwa, zwarcie linii, nadmierne tłumienie lub poziom szumu, wyłączony modem) a nie kwalifikacji linii. Możliwa też jest transmisja „Ethernet over ADSL”.

6. ZAKOŃCZENIE

Usługi POTS i ISDN w zasadzie osiągnęły nasycenie przy gęstości odpowiednio ok. 60% i 30% (w Polsce liczby te są z różnych przyczyn znacznie niższe) a czynnikiem powodującym rozwój i przekształcenie sieci dostępowych jest transmisja danych, związana z dostępem do globalnej sieci i innymi usługami (np. VPN). Dzięki postępowi technik kompresji obrazu i wzrostowi możliwości xDSL, po początkowych niepowodzeniach, większego znaczenia nabiera wideo na żądanie (VoD) a kolejną propozycją rozrywkową, stawiającą ostre wymagania sieci dostępowej są gry sieciowe.

Wzrost szybkości transmisji w nowych wersjach ADSL pozwala im na zastąpienie VDSL dla odległości rzędu 1 km (FTTC) a uzupełnieniem oferty w sieci miedzianej są modemy SHDSL, umożliwiające transmisję symetryczną ze zmienną szybkością.

Stanem, do którego dąży się obecnie, jest sieć dostępową określaną jako FSAN (Full Service Access Network), która będzie udostępniała pełny katalog usług, zarówno wąskopasmowych przy ujednoliconym wewnętrznym standardzie przesyłania informacji, dostosowanym do otoczenia telekomunikacyjnego i dominującego charakteru usług (transmisja pakietowa) [9], przy zachowaniu wymogów dotyczących mechanizmów zarządzania, QoS, billingu, zdalnego zasilania terminali i innych. Przejście od istniejących sieci dostępowych do FSAN może następować ewolucyjnie, co pozwala na wykorzystanie istniejących zasobów i dostosowywanie do zmieniających się potrzeb użytkowników. W perspektywie jest oferowanie usług przy pomocy terminali wyposażonych w interfejsy dostosowane do sieci IP, rozszerzając proces integracji na UNI i terminale abonenckie.

BIBLIOGRAFIA

- [1] Smoleński L., Zientalski M.: *Szerokopasmowe wielousługowe sieci dostępowe*, ZN WETI PG Tech. Inf. t.I, s. 337-346, Gdańsk 2003.
- [2] ITU-T: Rec.G.902, *Framework Recommendation on functional access networks*, 11/1995.
- [3] ITU-T: Rec.G.983.1, *Broadband optical access systems based on PON*, 10/1998.
- [4] ITU-T: Rec.G.983.3, *A broadband optical access system with increased service capabil.*, 03/2001
- [5] ITU-T: Rec.G.992.2, *Splitterless asymmetric digital subscriber line(ADSL) transceivers*, 06/1999.
- [6] Heron R., Van Parijs N.: *Evolutionary Pathways for the Broadband Access Network*, Alcatel Tel. Rev., 2/2003.
- [7] Crawford M., Verheye D.: *Residential Service Aggregation in the Second Mile*, AlcatelTR, 2/2003
- [8] *ADSL2 and ADSL2PLUS –The new ADSL standards*, White paper, www.dslforum.org, 03/2003.
- [9] *New Service Architectures for DSL Networks*, White Paper, www.alcatel.com, 11/2003.

EVOLUTION TRENDS OF TRANSMISSIONS TECHNIQUES IN MULTI-SERVICES OPTICAL AND COPPER ACCESS NETWORKS

Summary

This paper describes evolution trends in access networks, which used copper and fiber optics lines. These changes lead to obtain network that offered full set of broad- and narrowband services (FSAN). For the purpose of such network is to offer its users conventional and new services, with level of quality and reliability typical of telecommunications network. Changes in access networks consist in growing share of optical transmission, spreading of copper line asymmetric transmission rate systems (ADSL), aspiration to adjust transmission techniques and protocols with core and metropolitan networks (ATM, IP, optical layer). Successive evolutionary variants of access network drive at full integration, when all services would be realized in common service node.

Jacek Światowiak*, Krzysztof Nowicki**

*PSG Gdańsk, **Katedra Systemów Informacyjnych WETI Politechnika Gdańska

ZASADY WSPÓŁPRACY SIECI IPv4 I IPv6 ORAZ KIERUNKI MIGRACJI DO IPv6

Streszczenie

W artykule przedstawiono metody współpracy sieci IPv4 z sieciami pracującymi z protokołem IPv6. Omówiono możliwe warianty migracji z IPv4 do IPv6, w tym wykorzystanie translacji protokołów. Zwrócono uwagę na niebezpieczeństwo zahamowania rozwoju europejskiego Internetu spowodowane zbyt wolnymi procesami migracji do IPv6.

1. WSTĘP

Protokół IPv4, zaprojektowany ponad 20 lat temu [1], pomimo swoich niezaprzeczalnych zalet, posiada też sporo wad i ograniczeń, które utrudniają a wkrótce mogą uniemożliwić jego dalsze funkcjonowanie w nowoczesnych sieciach teleinformatycznych. Ograniczenia te wynikają zarówno ze struktury nagłówek warstwy sieciowej stosu TCP/IP wersja 4, jak też niewystarczających mechanizmów QoS, zapewnienia bezpieczeństwa czy też autokonfiguracji węzłów.

Nowa wersja protokołu IP (IPv6) [2], mimo, że powstała na bazie protokołu IPv4, zawiera wiele elementów i mechanizmów, które nie mają odpowiednika w IPv4. Stąd nie istnieje sposób prostej współpracy obu tych protokołów między sobą.

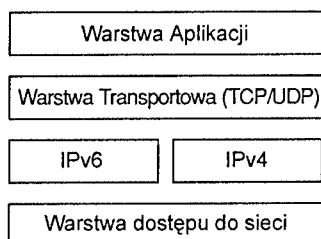
Realizacja wymiany informacji między węzłami IPv4 i IPv6 zależy od typów węzłów uczestniczących w komunikacji oraz konfiguracji połączeń. W wymianie informacji mogą brać udział węzły, które wspierają tylko protokół IPv4 (*tylko-IPv4* – ang. only IPv4), albo tylko protokół IPv6 (*tylko-IPv6* – ang. only IPv6), lub oba protokoły (węzły z podwójnym stosem IPv4/IPv6).

Węzły te mogą znajdować się na tym samym łączu albo w różnych sieciach. W przypadku komunikacji między węzłami znajdującymi się na tym samym łączu zarówno specyfikacja jak i implementacja komunikacji jest stosunkowo prosta. W przypadku węzłów IP pracujących w różnych sieciach komunikujących się przez sieć/sieci IPv4, implementacje komunikacji, w zależności od tego czy są to węzły typu *tylko-IPv6* czy też węzły zarówno typu IPv6 jak i IPv4, różnią się znacznie zarówno złożonością jak też zakresem wspieranych usług.

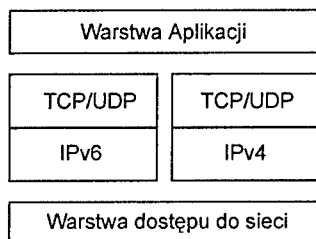
2. KOMUNIKACJA MIĘDZY WĘZŁAMI ZNAJDUJĄCYMI SIĘ NA TYM SAMYM ŁĄCZU

Dla węzłów IPv4/6 znajdujących się na tym samym łączu (przykładowo: w jednej domenie rozgłoszeniowej Ethernet), przynajmniej na jednym z pary węzłów musi zostać zainstalowany tzw. podwójny stos IPv4/IPv6, zapewniający komunikację zarówno z węzłami pracującymi z protokołem IPv4 jak i z IPv6.

Dwuprotokołowy stos IPv4/IPv6 może być realizowany w formie dual layer (rys. 1) lub dual stack (rys.2) [3].



Rys. 1 Architektura Dual Layer



Rys. 2. Architektura Dual Stack

Rozwiązanie dual stack (stosowane między innymi w systemie Microsoft Windows 2003) jest bardziej uniwersalne zapewniając większe bezpieczeństwo oraz łatwiejsze (niezależne od siebie) zarządzanie protokołami warstwy sieciowej i transportowej. Podwójny stos wykorzystywany jest również w realizacji tunelowania i translacji protokołów (patrz rozdz. 3 i 4)

Zgodnie z koncepcją warstwowej architektury sieciowej zarówno warstwa dostępu do sieci jak i aplikacji jest (powinna być) całkowicie niezależna od stosowanych protokołów TCP/IP. Niestety w wielu przypadkach tak nie jest.

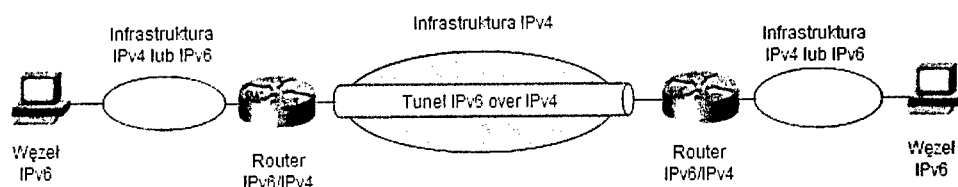
3. KOMUNIKACJA MIĘDZY WĘZŁAMI IPV6 ZNAJDUJĄCYMI SIĘ W RÓŻNYCH SIECIACH

W przypadku komunikacji węzłów IPv6 pracujących w różnych sieciach i komunikujących się ze sobą poprzez sieć/sieci IPv4 musi zostać zapewniony transport datagramów przez sieć/sieci IPv4. W tym celu stosowany jest mechanizm tunelowania IPv6 w IPv4.

3.1. Tunelowanie

Przy realizacji tunelowania datagramy IPv6, z punktu widzenia protokołu IPv4, traktowane są jako dane. Wymaga się jedynie umieszczenia w nagłówku datagramu IPv4 informacji o typie protokołu (wartość 41 identyfikuje protokół IPv6). Adres źródłowy i docelowy w datagramie IPv4 określają końce tunelu.

Wyróżniamy następujące rodzaje tuneli realizowane w sieciach IPv4: router-to-router (patrz rys.3), host-to router, host-to-host [4]. Dwa ostatnie rodzaje tuneli są jedynie uproszczeniem wersji routet-to-router.



Rys. 3. Tunelowanie typu router-to-router

W przypadku tunelowania typu router-router datagramy z węzła IPv6 w jednej sieci przesyłane są do routera dwuprotokołowego IPv6/IPv4, którego zadaniem jest enkapsulacja datagramu IPv6 w datagram IPv4 (IPv6->IPv4) oraz zestawienie tunelu do innego routera dwuprotokołowego IPv6/IPv4 znajdującego się w odległej sieci. Zadaniem tamtego routera jest dekapulacja IPv6 -> IPv4 i przesłanie danych do węzła końcowego.

Zasada działania tunelowania w dwóch pozostałych rodzajach tuneli jest podobna z tym, że funkcję enkapsulacji/dekapulacji IPv6 w IPv4 realizuje pojedynczy router lub sam host.

Tunele dzielimy na tunele konfigurowane manualnie przez administratorów sieci oraz tunele konfigurowane automatycznie z wykorzystaniem mechanizmów autokonfiguracji [5]. W przypadku manualnej konfiguracji tunelowania w celu zestawienia tunelu wymagane jest wyłączenie kilku poleceń na routerach/hostach. Obecnie wykorzystuje się następujące rozwiązania autokonfiguracji tuneli:

1. Autokonfiguracja z wykorzystaniem adresów IPv6 kompatybilnych z IPv4 [6],
2. 6to4 [7],
3. ISATAP (*ang. Intra-Site Automatic Tunnel Addressing Protocol*) [8],
4. 6over4 w przypadku możliwości zestawiania transmisji multicastowych [9],
5. Architektura TEREDO [10],
6. Architektura SILKROAD [11].

Pierwsze cztery nie wspierają sieci rozdzielonych urządzeniami NAT (*Network Address Translator* – [12])

3.1.1. Autokonfiguracje tuneli nie wspierających NAT

Autokonfiguracje tuneli 6to4, 6over4 oraz wykorzystujących adresy IPv6 kompatybilne z IPv4 zwalniają administratorów z konieczności manualnej konfiguracji tuneli. Posiadają one niestety jedną aczkolwiek dużą wadę – wymagają, aby urządzenia zestawiające tunel z obu stron posiadały globalny/publiczny adres IPv4. Wewnątrz wielu sieci, zarówno korporacyjnych jak i prywatnych, wykorzystuje się często adresy z pól prywatnych. Powyższe rozwiązania są wówczas nieprzydatne.

Z kolei rozwiązanie ISATAP będące pochodną rozwiązania 6to4 wykorzystuje do zestawiania tuneli routery 6to4 pracujące na brzegu sieci (routery brzegowe), zaś wewnątrz sieci może posługiwać się adresami z pul prywatnych. Umożliwia to zestawianie tuneli nawet w przypadku, gdy obie podsieci prywatne wykorzystują identyczne (niepubliczne) pule adresowe. Niestety, rozwiązanie ma identyczne wady jak 6to4. Na brzegu sieci musimy mieć router 6to4, co niestety bardzo często jest bardzo trudne do realizacji. Obecnie większość sprzętu sieciowego dla sieci WAN (poza urządzeniami CISCO oraz routerami programowalnymi – typowe rozwiązania „open source” – LINUX, BSD) nie ma zaimplementowanego ani tunelowania 6to4 ani rozwiązania ISATAP.

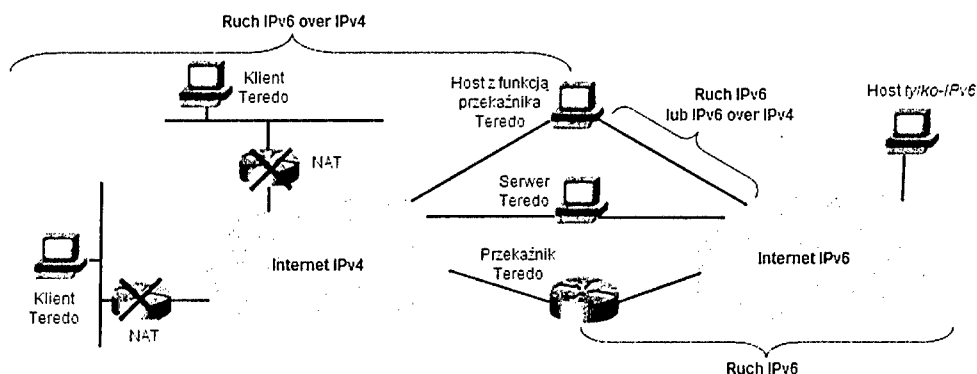
Tunelowanie 6over4 jest obecnie rzadko stosowane, gdyż może być realizowane jedynie w rozwiązaniach sieciowych wspierających transmisję multicastową, co poza sieciami LAN (w szczególności Ethernet) wymagałoby emulacji transmisji multicastowej na łączach typu NBMA (ang. *Non-Broadcast Multi Access*) jak w ATM czy Frame-Relay. Rozwiązanie to jest byłoby zatem kosztowne i trudne do realizacji.

3.1.2. Autokonfiguracje tuneli wspierające NAT

Większość problemów opisanych powyżej zostało rozwiązanych w projektach TEREDO [10] oraz jego następcy – SILKROAD [11].

3.1.2.1. TEREDO

TEREDO przede wszystkim umożliwia automatyczne adresowanie węzłów oraz tworzenie tuneli typu węzeł-węzeł dla unicastowej komunikacji pomiędzy węzłami IPv6/IPv4 znajdującymi się za jednym bądź wieloma urządzeniami IPv4 NAT. Aby przejść przez urządzenia IPv4 NAT, pakiety IPv6 są wysyłane jako komunikaty IPv4 UDP.



Rys. 4. Komponenty architektury TEREDO

W skład architektury TEREDO wchodzi następujące elementy: klient TEREDO, serwer TEREDO, przełącznik TEREDO oraz klient z funkcją przełącznika (patrz rys.4).

TEREDO pomimo swoich zalet (współpraca z urządzeniami NAT wielu producentów, brak konieczności modyfikacji istniejących już urządzeń NAT, przezroczystość infrastruktury IPv4 dla rozwiązania TEREDO) nie jest wolna od wad. Obecnie są wykorzystywane cztery typy urządzeń NAT [13]: *cone*, *restricted cone*, *port restricted cone* i *symmetric*. Niestety, TEREDO nie ze wszystkimi urządzeniami NAT może prawidłowo współpracować lub wymaga ingerencji administratora (np. przekierowania portów).

Powyższe rozwiązania posiadają wiele ograniczeń, w tym: nie rozwiązane do końca kwestie bezpieczeństwa każdego z rozwiązań, brak wsparcia QoS, ograniczoną wydajność oraz trudności w administracji.

3.1.2.2. SILKROAD

Najnowszym rozwiązaniem, pozbawionym wad TEREDO, jest SILKROAD. Podobnie jak TEREDO wykorzystuje enkapsulację datagramów IPv6 w pakietach UDP IPv4. Jednakże umożliwia realizację komunikacji IPv6 przez wszystkie typy urządzeń NAT, w tym *symmetric NAT*. Nie wymaga przy tym żadnego specjalnego prefixu.

SILKROAD będzie mógł być również wykorzystywany w przyszłości, gdy węzły IPv4 będą tylko małymi wyspami w Internecie IPv6.

4. KOMUNIKACJA POMIĘDZY WĘZŁAMI *tylko-IPv4* I *tylko-IPv6* ZNAJDUJĄCYCH SIĘ W JEDNEJ LUB WIELU SIECIACH

Odmianą od tunelowania metodą zapewniania komunikacji pomiędzy węzłami IPv4 i IPv6 jest dwukierunkowa translacja protokołów IPv4 <-> IPv6. Nie dotyczy ona wyłącznie warstwy sieciowej, ale również transportowej oraz aplikacji (w szczególności protokołów ICMP, TCP, UDP oraz np. WWW).

Technologia translacji pomiędzy protokołami IPv4 a IPv6 jest różna w zależności czy dokonujemy translacji z IPv4 do IPv6 czy też z IPv6 do IPv4.

Do podstawowych rozwiązań zaproponowanych przez organizację IETF należą:

- SIIT (ang. *Stateless IP/ICMP Translation Algorithm*) [14]
- NAT-PT (ang. *Network Address Translation – Protocol Translation*) oraz NAPT-PT (ang. *Network Address Port Translation + Protocol Translation*) [15],
- BIS (ang. *Bump In the Stack*) [16],
- BIA (ang. *Bump In the API*) [17],
- TRT (ang. *Transport Relay Translator*) [18],
- ALG (ang. *Application Layer Gateway*) [15, 19],
- SOCKS IPv4/IPv6 [20].

4.1. SIIT

Ze względu na istotne różnice pomiędzy protokołami IPv4 a IPv6 (odmienne struktury datagramów i nagłówków, w tym przeniesienie części opcji z nagłówka IPv4 do nagłówków dodatkowych IPv6, zmiany zasad współpracy z warstwą dostępu do sieci) występuje wiele problemów przy konwersji pomiędzy nimi.

Część parametrów zawartych w nagłówkach (numery wersji, typ-usługi (priorytety), czas życia (limit skoków)) można dość łatwo przekonwertować. Jednakże bezpośrednia konwersja np. adresu IPv4 na adres IPv6, adresu IPv6 na adres IPv4 nie jest możliwa, gdyż pola adresowe mają różną długość (32 bity w IPv4 i 128 bitów w IPv6). Niemożliwa jest również konwersja niektórych pól, gdyż występują one tylko w jednej wersji protokołu (np. etykieta przepływu).

Dodatkowym problemem, w warstwie aplikacji, jest odwzorowanie adresów IP na nazwy stacji roboczych czy serwerów – dotyczy to usługi DNS i konieczności obsługi dodatkowych rekordów związanych z IPv6: AAAA oraz A6.

SIIT opisuje mechanizm translacji protokołów pomiędzy węzłami *tylko-IPv6* i *tylko-IPv4* polegający na translacji informacji zawartych w nagłówkach datagramów IPv4 i IPv6 bez konieczności wykorzystywania informacji o aktualnie realizowanych protokołach warstwy transportowej. Każde połączenie pomiędzy dwoma węzłami IPv4 <-> IPv6 poprzez translator SIIT wykorzystuje jeden publiczny adres z dostępnej puli. Poza koniecznością posiadania publicznych adresów IPv4 mechanizm SIIT posiada inną dość poważną wadę: nie jest w stanie zapewnić autoryzacji IPv6 z wykorzystaniem nagłówka AH (ang. *Authentication Header*). Stanowi to poważny problem w przypadku wykorzystywania protokołu IPSec.

4.2. NAT-PT, NATP-TP

Protokół NAT-PT do translacji datagramów IPv4 <-> IPv6 wykorzystuje te same zasady, które zdefiniował SIIT. NAT-PT, podobnie jak SIIT, wykorzystuje pulę globalnych adresów IPv4 w celu przyporządkowania ich do węzłów IPv6, jednakże w sposób dynamiczny – tylko na czas jednej sesji połączenia. Urządzenia (routery) NAT-PT zapewniają transparentny routing datagramów pomiędzy siecią IPv4 a IPv6. Technologia ta nie wymaga żadnych zmian w węzłach sieci czy to IPv4 czy IPv6. Wymaga jednak śledzenia przez urządzenia NAT-PT sesji tak, aby datagramy przechodziły przez określony router NAT-PT.

Protokoły z rodziny NAT-PT można podzielić na: tradycyjny (jednokierunkowy), gdzie sesje są inicjowane przez węzły IPv6 z urządzenia NAT-PT do węzłów IPv4, oraz dwukierunkowy, gdzie oba typy węzłów mogą inicjować transmisję. Tradycyjny NAT-PT można dodatkowo podzielić na tzw. podstawowy (Basic-NAT-PT) oraz na NAPT-PT. Basic-NAT-PT dokonuje jedynie translacji adresów IPv6 <-> IPv4 oraz koniecznych zmian w protokołach pokrewnych jak ICMP, TCP i UDP w tym i odpowiednich sum kontrolnych.

NAPT-PT rozszerza funkcjonalność Basic-NAT-PT o możliwość zmiany numerów portów TCP, UDP oraz identyfikatorów komunikatów ICMP. Umożliwia to wielu węzłom IPv6 współdzielenie w urządzeniach NAPT-PT jednego publicznego adresu IPv4 w wielu sesjach.

W dwukierunkowym NAT-PT każda ze stron może nawiązać połączenie. Adresy IPv4 wykorzystywane w urządzeniu NAT-PT mogą być przydzielane na czas połączenia statycznie bądź dynamicznie. Aby zapewnić dostęp węzłów IPv4 do IPv6 wykorzystuje się usługę DNS w celu rozwiązywania nazw. Jednakże DNS wymaga wsparcia usługi ALG, która tłumaczy adresy IPv6 w zapytaniach DNS na przydzielone im adresy IPv4 oraz komunikaty DNS pomiędzy sieciami IPv4 a IPv6.

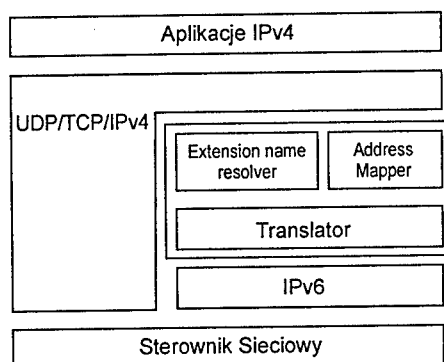
Dwukierunkowy NAT-PT rozszerzony o NAPT-TP mogłyby być dominującym rozwiązaniem translacji datagramów IPv6 <-> IPv4, gdyby nie konieczność wykorzystywania ALG dla wielu operacji protokołowych jak DNS, FTP itd.

Ważnym problemem jest również zapewnienie bezpieczeństwa typu koniec-koniec. Protokół IPsec nie może być wykorzystywany przy przejściu przez urządzenia NAT-PT. Dodatkowo w przypadku DNS-ALG nie jest możliwe wykorzystywanie Secure DNS, co mogłoby być wykorzystane przy ataku typu DNS-SPOOFING.

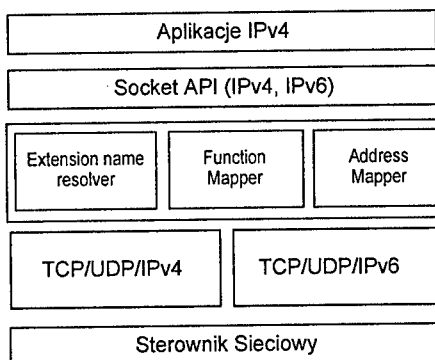
4.3. BIS i BIA

O ile rozwiązania NAT-PT, NAT-TP i SIIT miały na celu tłumaczenie protokołów, to celem BIS i BIA jest zapewnienie kompatybilności aplikacji użytkownika zaprojektowanych dla IPv4 z protokołem IPv6.

W architekturze BIS dodano do standardowego stosu TCP/IPv4 dodatkowy moduł, który składa się z stosu IPv6, translatora, address mappera i extension name resolvera (patrz rys.6). Zadaniem tego modułu jest translacja protokołu IPv4 <-> IPv6. Podczas komunikacji z węzłem IPv6 tworzony jest, na podstawie adresu IPv4, adres IPv6 i przydzielany do interfejsu. Dzięki działającej usłudze DNS, użytkownik nie musi wiedzieć czy komunikuje się z węzłem IPv4 czy IPv6. Węzeł taki z punktu widzenia aplikacji zachowuje się jakby posiadał całkowicie niezależny podwójny stos.



Rys. 6. Architektura stosu BIS



Rys. 7. Architektura stosu BIA

Niestety, w przypadku aplikacji wykorzystujących specyficzne właściwości obu protokołów, nie będzie możliwe przeprowadzenie konwersji między nimi. Dostępne będą tylko elementy związane z fragmentacją (*IPv6 Fragment Header*) i routowaniem (*IPv6 Routing Header*).

Dodatkowym problemem jest dokonywanie konwersji pomiędzy adresami IP zawartymi wewnątrz warstwy aplikacji, np. ftp, oraz protokołem Secure DNS, gdyż extension name resolver nie wspiera tego protokołu. Metoda ta wymaga dużej integracji ze stosowanym sterownikiem interfejsu sieciowego.

Mechanizm BIS będzie działał poprawnie w przypadku transmisji unicastowej. Natomiast dla transmisji multicastowej należy stosować inne mechanizmy.

Z kolei w BIA translator API (ang. *Application Programming Interfaces*) umieszczony jest pomiędzy socket API a podwójnym stosem TCP/IP. Umożliwia to tłumaczenie IPv4 socket API w IPv6 socket API i odwrotnie. W takim przypadku translacja protokołów IPv6 jest prostsza, odbywa się bez konieczności translacji nagłówków IPv6 i IPv4 i z wykorzystaniem podwójnego stosu typu dual stack (patrz rys.7). Gdy warstwa aplikacji IPv4 nawiązuje komunikację się z węzłem IPv6, translator rozpoznaje funkcje socket API pochodzące od protokołu IPv4 i wywołuje funkcje IPv6 socket API w celu nawiązania komunikacji z węzłem IPv6 i odwrotnie.

Podobnie jak w NAT-PT, BIA nie jest w stanie dokonywać translacji adresów IP zawartych w warstwie aplikacji, np. ftp. Zatem nie będzie on współpracować z tego typu aplikacjami. Podobnie jak BIS będzie działał poprawnie tylko w przypadku transmisji unicastowej.

4.5. SOCKS IPv4/IPv6

SOCKS – popularne „w czasach systemów klasy Windows NT 4.0” rozwiązanie zapewniające dostęp aplikacjom użytkownika do aplikacji znajdujących się na innych stacjach roboczych, często w innych sieciach.

Mechanizm bramy SOCKS IPv6/IPv4 realizuje przez te urządzenia dwie niezależne transmisje IPv4 and IPv6 na poziomie warstwy aplikacji. Translacja bazująca na bramie SOCKS IPv6/IPv4 jest dość trudna do implementacji, gdyż wymaga poza serwerem SOCKS, instalacji w węźle klienta dodatkowych bibliotek (SOCKS API).

4.6. TRT i ALG

IPv6-to-IPv4 Transport Relay Translator (TRT) i ALG (Application Layer Gateway) nie definiują żadnego nowego rodzaju translacji protokołów. Są one raczej sposobem implementacji istniejących rozwiązań translatorów i nie wnoszą żadnych nowych rozwiązań ani protokolarnych ani implementacyjnych.

5. KIERUNKI MIGRACJI DO IPV6

Problem migracji nie leży ani w ograniczeniach sprzętowych ani w programowych, gdyż nie istnieje żadne technologiczne ograniczenie w budowie sieci opartych o IPv6 (IETF zdefiniowało sposoby transmisji datagramów we wszystkich istniejących typach sieci [12]) oraz jest pełna dostępność implementacji IPv6 w każdym sieciowym systemie operacyjnym. Okazuje się, iż problemem jest „podejście decydentów”, administratorów sieci szkieletowych, tych dużych i mniejszych, dostawców usług internetowych, integratorów sieci i systemów informatycznych, producentów sprzętu a głównie oprogramowania w kwestii wdrażania protokołu IPv6.

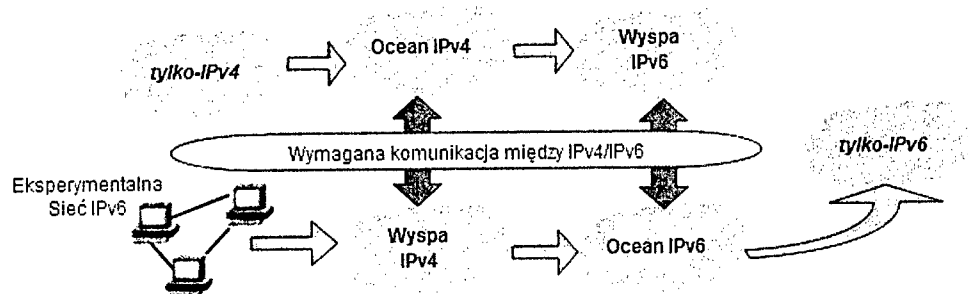
5.1. Scenariusze migracji

Mimo teoretycznej niezależności wszystkich z warstw stosu TCP/IP, zależności istnieją i sięgają od warstwy dostępu do sieci aż do aplikacyjnej. Niemożliwa jest zatem prosta translacja protokołów IPv4 <-> IPv6. Wymaga to, niestety, wielu zmian w oprogramowania zarówno użytkownika jak i systemowym.

Odpowiadając na pytanie jakich scenariuszy migracji możemy się spodziewać, przewiduje się następujące fazy:

- powstawanie „wysp” IPv6 w „oceanie” IPv4. Sieci IPv6, które będą się komunikować z innymi sieciami IPv6 poprzez infrastrukturę IPv4 – metodą tunelowania, tak jak aktualnie realizuje się korporacyjne sieci VPN.
- dostęp węzłów IPv4/IPv6 znajdujących się w prywatnych podsięciach (niepubliczne adresy IP) za pomocą technologii TEREBO lub SILKROAD.
- migracja sieci szkieletowych (ATM/MPLS/POS/GETH) i dostępowych (Frame-Relay/xDSL) do dwuprotokołowych IPv4/IPv6. Zapewni to lepsze wykorzystanie możliwości w dziedzinie QoS czy bezpieczeństwa, które oferuje IPv6.
- migracja sieci szkieletowych do czystego IPv6. Protokół IPv4 będzie tunelowany w IPv6,
- zanik „Internetu IPv4” i przejście do „Internetu IPv6” z „wyspami IPv4”, które to zawierają węzły IPv4, które nie mogą zostać przekształcone do IPv6,
- docelowo całkowity zanik protokołu IPv4.

Fazy migracji z sieci *tylko-IPv4* do sieci *tylko-IPv6* ilustruje rys. 8.



Rys. 8. Kierunki migracji z IPv4 do IPv6

5.2. Praktyczny przykład procesu migracji

Wśród wielu zrealizowanych przykładów migracji IPv4 do IPv6 (IPv6 over MPLS, IPv6 over ATM, IPv6 over Gigabit Ethernet) niezwykle interesującym (i zaawansowanym technologicznie) dla sieci szkieletowych jest rozwiązanie IPv6 over MPLS realizowane w trzech trybach:

1. Native IPv6 over MPLS [21] – transport datagramów IPv6 po sieci MPLS wymaga jednak aby wszystkie routery MPLS były dwu-protokołowe oraz współpracowały z protokołami routingu dla IPv6 (wewnętrznego i zewnętrznego) oraz z protokołem LDP (ang. Label Distribution Protocol).
2. Layer 2 tunneling over MPLS [22] – kompletne ramki warstwy drugiej (Ethernet wraz z rozszerzeniem IEEE 802.1q, ATM AAL5) są przełączane/dystrybuowane bezpośrednio po sieci MPLS stąd sieć ta jest całkowicie przezroczysta dla warstw wyższych. Opcje te są już dostępne np. w sprzęcie CISCO i Juniper.
3. IPv6overIPv4/MPLS Core [23] – metoda dystrybucji prefixów IPv6 (i odpowiadającym im etykiatom MPLS) po ścieżce LSR (ang. *Label-Switching Routers*) z wykorzystaniem standardowego protokołu BGP4+. Opcja ta zaimplementowana jest np. w routerach CISCO pod nazwą 6PE (IPv6 Provider Edge Router)

6. PODSUMOWANIE

Pomimo zdefiniowania już wielu metod translacji protokołów IPv4 <-> IPv6 doświadczenia autorów niniejszej pracy pozwalają stwierdzić, iż technologicznie łatwiejsze i znacznie mniej kosztowne będzie w procesie migracji do IPv6, wykorzystywanie technik tunelowania zamiast translacji protokołów. Pomimo wielu ograniczeń, jakie niesie za sobą tunelowanie (w tym brak wspierania QoS) istotne jest, że jego funkcjonowanie zostało praktycznie przetestowane i jest realizowane w wielu środowiskach sieciowych [12].

Zrealizowane są już projekty i wdrożenia sieci IP bazujące na obu protokołach jak i sieci *tylko-IPv6* (6 bone [24], 6 ren [25], GEANT [26], moonv6 [27]).

W Europie procesy migracji przebiegają wolno. Autorzy nie rozważali w artykule aspektów natury politycznej, prawnej czy ekonomicznej, które determinują przyjęcie i wykorzystywanie takich a nie innych rozwiązań technologicznych. Jednak wydaje się, iż obecnie migracja do IPv6 wymaga w Europie jedynie poparcia politycznego. Warto zauważyć, że w USA agendy rządowe, w tym armia, rozpoczęły już dawno proces migracji do IPv6 – termin całkowitego przejścia na IPv6 wyznaczono na rok 2007, a od roku 2004

wszystkie nowe urządzenia muszą wspierać IPv6. Azja zmuszona została do migracji z powodu braku dostępnych publicznych adresów IPv4.

Najgorszym scenariuszem dla Europy (więc i Polski) może okazać się stagnacja w kwestiach decyzyjnych, która to może spowodować zapóźnienie technologiczne i w konsekwencji coraz większe problemy i wyższe koszty komunikacji pomiędzy Europą a resztą Świata.

LITERATURA

- [1] RFC 791 Internet Protocol DARPA Internet Program Protocol Specification
- [2] RFC 2460 Internet Protocol, Version 6 (IPv6) Specification
- [3] IETF draft Dual Stack Transition Mechanism (DSTM)
- [4] RFC 2893 Transition Mechanisms for IPv6 Hosts and Routers
- [5] RFC 3056 Connection of IPv6 Domains via IPv4 Clouds
- [6] RFC 3484 Default Address Selection for Internet Protocol version 6 (IPv6)
- [7] RFC 2893 Transition Mechanisms for IPv6 Hosts and Routers
- [8] IETF draft draft-ietf-ngtrans-isatap-0x.txt
- [9] RFC 2529 Transmission of IPv6 over IPv4 Domains without Explicit Tunnels
- [10] IETF draft Tunneling IPv6 over UDP through NATs,
- [11] IETF draft SILKROAD – „Tunneling IPv6 with private IPv4 addresses behind NAT devices”
- [12] Nowicki K., Światowiak J.: Protokoły IPv6. Opis protokołów. Materiały do laboratorium. Wydawnictwo PG 2001.
- [13] RFC 3489 STUN – Simple Traversal of User Datagram Protocol(UDP)Through Network Address Translators (NATs)
- [14] RFC 2765 Stateless IP/ICMP Translation Algorithm (SIIT)
- [15] RFC 2663 IP Network Address Translator (NAT) Terminology and Considerations
- [16] RFC 2767 Dual Stack Hosts using the "Bump-In-the-Stack" Technique
- [17] RFC 3338 Dual Stack Hosts using Bump-in-the-API
- [18] RFC 3142 An IPv6-to-IPv4 Transport
- [19] RFC 2663 IP Network Address Translator (NAT) Terminology and Considerations
- [20] RFC 3089 A SOCKS-based IPv6/IPv4 Gateway Mechanism
- [21] IETF draft BGP-MPLS VPN extension for IPv6 VPN
- [22] IETF „Encapsulation Method of Transport of Layer 2 Frames Over IP and MPLS Networks
- [23] IETF draft „Connectiong IPv6 Islands cross IPv4 Clouds with BGP
- [24] 6bone – <http://www.6bone.net>
- [25] 6ren – <http://www.6ren.net>
- [26] GEANT – Multi-Gigabit pan-European Research Network, <http://www.dante.net>
- [27] Project moonv6, <http://moonv6.sr.unh.edu/>

INTERNETWORKING BETWEEN IPV4 AND IPV6 PROTOCOLS. DIRECTIONS OF MIGRATION TO IPV6

Abstract

The article presents protocol's solution of cooperation in networks with IPv4 and IPv6 protocols. It describes possible variants of migration from IPv4 to IPv6 and ways of translations the protocols between IPv4 and IPv6. The authors pay attention to dangerous slow down of expansion in European Internet, made by too slow decision process of migration to IPv6.

Liliana Byczkowska-Lipińska, Marcin Cegielski

Instytut Informatyki, Politechnika Łódzka

ARCHITEKTURY SYSTEMU INFORMATYCZNEGO DLA PRZEDSIĘBIORSTW MAŁEJ I ŚREDNIEJ WIELKOŚCI

Streszczenie

Potrzeby, jakie są stawiane w dobie gospodarki elektronicznej przed firmą małą lub średniej wielkości muszą być rozwiązane przez system informatyczny. Rozwój telekomunikacji powoduje znaczący wzrost możliwości wymiany informacji pomiędzy niezależnymi podmiotami gospodarczymi. Usługi sieciowe (Web Services), XML, WSDL, SOAP, DCOM pozwalają na organizację szybkiej wymiany danych między różnymi systemami informatycznymi. Pozwala to na przejęcie przez system wielu czynności wykonywanych do tej pory przez człowieka. System informatyczny wymaga pracy na wielu platformach sprzętowych. W artykule zostaną omówione narzędzia potrzebne do stworzenia takiego systemu dla małych i średnich przedsiębiorstw, który spełnia powyższe wymagania. Podane zostaną koncepcje i kierunki rozwoju systemów środowisk programistycznych wykorzystywanych do tworzenia i budowy systemów informatycznych.

1. WPROWADZENIE

Opracowania teoretyczne omawiające wdrażanie rozbudowanych systemów informatycznych mają charakter organizacyjny lub ekonomiczny i dotyczą przedsiębiorstw dużych lub korporacji. Małe i średnie firmy ze względu na możliwości własnego budżetu przeznaczanego na informatyzację praktycznie były pozbawione implementacji zaawansowanych technologii informatycznych. Rozwój informatyki, usług telekomunikacyjnych i ich coraz większa powszechność pozwoliły na wdrożenia informatyczne dotąd zarezerwowane dla dużych przedsiębiorstw. Ponieważ środki finansowe w małych i średnich firmach są ograniczone, narzuca to konieczność opracowania takich systemów, aby koszt ich budowy mógł być rozłożony na kilka firm. Wymaga to konieczności sklasyfikowania przedsiębiorstw i zbudowania pewnych elementów takiego systemu, aby można było je implementować wiele razy.

2. OKREŚLENIE ZADAŃ REALIZOWANYCH PRZEZ SYSTEM INFORMATYCZNY

Podstawowym zadaniem systemów informatycznych w przedsiębiorstwach jest automatyzacja możliwie jak największej liczby procesów biznesowych przeprowadzanych w firmie. Kolejnym zadaniem jest integracja wielu programów i systemów w jeden spójny system. Rozwój i dostępność usług telekomunikacyjnych pozwala na rozproszoną pracę wielu oddziałów i wymianę danych w czasie rzeczywistym.

3. CZYNNIKI WPLYWAJĄCE NA REALIZACJĘ PROJEKTU INFORMATYCZNEGO

Większość firm posiada już pewną infrastrukturę informatyczną i potrzebne zasoby sprzętowe zapewniające pracę oprogramowania. Jeżeli istniejąca baza sprzętowa spełnia wymagania systemu informatycznego pozwalając jednocześnie na jej rozbudowę i zwiększenie wydajności w przyszłości, to koncepcja stworzenia systemu informatycznego powinna być oparta o już istniejącą bazę sprzętową. W praktyce ze względu na koszty w wielu przypadkach posiadane zasoby sprzętowe narzucają lub wręcz wymuszają zastosowanie odpowiedniego rozwiązania informatycznego.

Wymiana serwerów wraz z systemami operacyjnymi i aplikacjami baz danych jest bardzo kosztowna i stanowi znaczący udział w kosztach całej operacji informatyzacji. Dlatego prawidłowe zaplanowanie użycia narzędzi programistycznych, projektu i etapów jego wdrażania powinno uwzględniać jak największe wykorzystanie posiadanych środków i doświadczenia pracowników w danym przedsiębiorstwie. Środowisko systemów operacyjnych ich lokalizacja, infrastruktura telekomunikacyjna, a także charakter pracy i zadania przedsiębiorstwa mają podstawowy wpływ na dobór narzędzi programistycznych. Przy wyborze oprogramowania pomocna jest klasyfikacja przedsiębiorstw według, określonych poniżej zasad.

1. Struktury organizacyjnej:

- firma posiada jedną siedzibę główną,
- firma posiada siedzibę główną i kilka oddziałów terenowych,
- firma posiada kilka jednostek w pełni niezależnych samodzielnych (mogą wykonywać te same czynności).

2. Posiadanych i używanych łącz telekomunikacyjnych:

- wykorzystywanie łącz dzierżawionych lub kanałów VPN,
- wykorzystywanie publicznych stałych łącz internetowych we wszystkich oddziałach,
- wykorzystywanie stałych łącz internetowych w siedzibie głównej i łącz komutowanych w swoich oddziałach.

Ograniczenia, jakie występują przy doborze usług sieciowych są dwójakiego rodzaju.

- a) Lokalny stan infrastruktury telekomunikacyjnej, która pozwala na instalację w danej lokalizacji.
- b) Koszt samych usług telekomunikacyjnych, powinien być optymalny dla małego lub średniego przedsiębiorstwa.

Dla dużej firmy lub korporacji koszty eksploatacyjne łącz telekomunikacyjnych są pomijalnie małe, ale dla firmy średniej (25-50 pracowników) lub małej (poniżej 25 pracowników), są już liczącym się w koszcie w budżecie firmy.

Niestabilna sytuacja gospodarcza wielu małych i średnich firm pogłębia jeszcze problem ekonomiczny stałych kosztów związanych z zapewnieniem połączeń telekomunikacyjnych.

3. Używanej platformy sprzętowej i oprogramowania systemowego.

- firma używa tylko serwerów i stacji roboczych pracujących pod jednym systemem operacyjnym,
- firma używa serwerów i stacji roboczych pracujących pod wieloma systemami operacyjnymi,
- firma używa serwerów, stacji roboczych i urządzeń mobilnych pracujących pod wieloma systemami operacyjnymi. Przez urządzenia mobilne rozumie się Palmtopy i telefony komórkowe, których architektura sprzętowa znacznie się różni od komputerów.

4. Rodzaju posiadanych serwerów WWW i pocztowych:

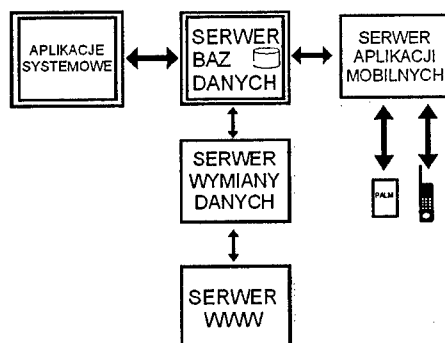
- serwer www i serwer pocztowy znajdują się w firmie,
- serwer www znajduje się w firmie, serwer pocztowy jest dzierżawiony w firmie trzeciej,
- serwery www i pocztowy są dzierżawione w firmie trzeciej.

Umieszczenie serwerów internetowych może całkowicie zdecydować o sprawności i możliwościach przyszłego systemu informatycznego.

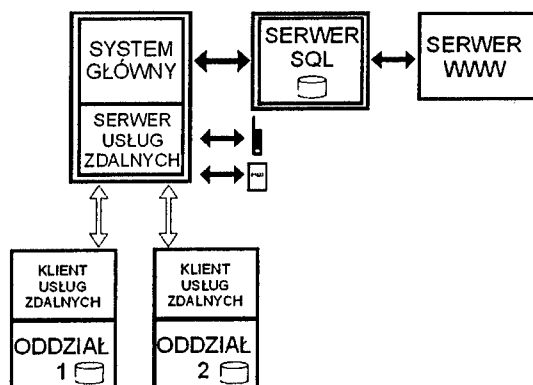
Ustalenie danych początkowych systemu powinno stanowić podstawę przy doborze odpowiednich narzędzi programistycznych, projektowych oraz może wpłynąć na architekturę przyszłego systemu informatycznego. Ustalenie założeń systemowych i struktury statystycznego przedsiębiorstwa handlowo-usługowego pozwala na zbudowanie kilku schematów przepływu danych w przedsiębiorstwie.

4. SCHEMAT PRZEPŁYWU DANYCH W SYSTEMIE

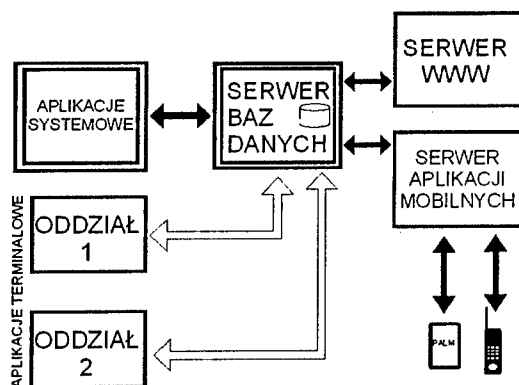
Prawidłowe zaprojektowanie przepływu danych w przedsiębiorstwie to już połowa sukcesu. Można ogólnie w prosty sposób przedstawić wymianę danych pomiędzy poszczególnymi elementami systemu informatycznego w przedsiębiorstwie. Na rys. 1,2,3 zostały przedstawione diagramy przepływu danych w zależności od zastosowanej infrastruktury informatycznej.



Rys. 1. Schemat wymiany danych w przedsiębiorstwie jedno-oddziałowym, posiadającym stałe łącza internetowe i dzierżawiącego serwer www.



Rys. 2. Schemat wymiany danych w przedsiębiorstwie wielo-oddziałowym, posiadającym stałe łącza internetowe i własny serwer www w siedzibie głównej i komutowany dostęp do internetu w oddziałach.



Rys. 3. Schemat wymiany danych w przedsiębiorstwie wielo-oddziałowym, posiadającym stałe łącza internetowe we wszystkich oddziałach i posiadające swój własny serwer www i pocztowy.

System wymiany danych pomiędzy niezależnymi firmami, zarządzanie łańcuchem dostaw i usługi sieciowe otwierają nowe możliwości do realizacji automatycznego przepływu danych pomiędzy współpracującymi ze sobą podmiotami gospodarczymi. Wiele przedsiębiorstw wdraża lub planuje wdrożenie systemu wymiany danych, ponieważ przynoszą one wymierne korzyści ekonomiczne, zwiększając możliwości planowania produkcji i sprzedaży. Przedsiębiorstwo posiadające wdrożone usługi elektronicznej wymiany dokumentów narzuca stosowanie ich swoim odbiorcom.

Rozwój informatyzacji zmusza małe i średnie firmy do rozbudowy systemów informatycznych i konieczność komunikowania się z innymi systemami. Najczęściej jest to realizowane za pomocą usług sieciowych (Web Services). Dane w postaci zbiorów XML przekazywane pomiędzy systemami mogą być dowolnie przeformatowywane. Przykładem takiego systemu jest elektroniczna wymiana dokumentów *ECOD* firmy *ComArch S. A.*, która pozwala na złożenie zamówienia, potwierdzenie zamówienia i przesłanie faktury

kończącej całą operację. System informatyczny przedsiębiorstwa może komunikować się ze wszystkimi usługami zarówno w trybie on-line jak i off-line. Możliwość korzystania z obu trybów pracy stwarza duże możliwości implementacji usług wymiany dokumentów w firmach o bardzo słabej infrastrukturze telekomunikacyjnej.

Stan i możliwości infrastruktury telekomunikacyjnej, jakie są do dyspozycji w przedsiębiorstwie, mają bezpośredni wpływ na zastosowanie odpowiednich narzędzi informatycznych takich jak usługi sieciowe i bazy danych. Rodzaj wybranej komunikacji pomiędzy oddziałami firmy może narzucić konieczność korzystania z usług aplikacji baz danych takich, jak replikacja czy system podwójnych transakcji. Podwójne transakcje w systemach baz danych to mechanizm podwójnego zatwierdzania transakcji przydatny w rozwiązywaniu problemów związanych ze spójnością danych w systemach rozproszonych.

5. IMPLEMENTACJA USŁUG B2B I B2C W PRZEDSIĘBIORSTWIE

Do znanych usług zapewniających komunikację z kontrahentami i użytkownikami należy zaliczyć systemy klasy B2B i B2C. Usługi B2B (*Bussiness To Bussiness*) i B2C (*Bussines To Customer*) okazują się niezbędnymi elementami systemów informatycznych przedsiębiorstw handlowych i handlowo-usługowych. Aplikacje te pozwalają obsługiwać jednocześnie bardzo wielu odbiorców w jednym czasie. Klienci danej firmy otrzymują dostęp do informacji handlowej w sposób czasowo nieograniczony przy jednoczesnym minimalnym zaangażowaniu pracowników danej firmy.

Aplikacja klasy B2B to system zorientowany na komunikację i obsługę odbiorców za pomocą usług www. Zapewnia ona wymianę informacji biznesowych pomiędzy firmą a jej kooperantami. Aplikacja klasy B2C jest przewidziana do komunikacji pomiędzy firmą a jej końcowymi odbiorcami

Informacja, przekazywana przez system informatyczny, powinna być pełna i aktualna. Jeżeli udostępnia się informację o stanach magazynowych, to klienci oczekują, że będą one cały czas aktualne. Aby to zapewnić należy połączyć usługi serwera internetowego www z bazami danych podstawowego systemu informatycznego. Koszty utrzymania aplikacji internetowej zależą głównie od architektury systemu informatycznego.

System klasy B2B musi być tak zbudowany, aby modyfikacja wszelkich danych biznesowych (oferta handlowa, promocje, dodatkowe informacje), była możliwa za pomocą tej samej aplikacji, którą pracownicy użytkują codziennie w pracy bez konieczności ingerencji wykwalifikowanego personelu informatycznego.

Systemy klasy B2B i B2C muszą zapewniać pełną integrację aplikacji internetowych ze wszystkimi systemami informatycznymi funkcjonującymi w danym przedsiębiorstwie. Przykładem może być zastosowanie serwera *IIS (Internet Inforation Server)*, *Share Poit Server* i pakietu *Office 2003* firmy *Microsoft*. Za pomocą tych aplikacji (*IIS*, *Share Poit Server*, *Office 2003*) pracownik opracowuje nową ofertę handlową i zapisuje ją w odpowiednim folderze. Każdy dokument, z którego informacja zostaje przekazana do systemu B2B lub B2C zostaje automatycznie przeformatowana do postaci html i szaty graficznej zgodnej z obowiązującą formą na stronach www.

6. ARCHITEKTURA SYSTEMU INFORMATYCZNEGO

Jeżeli przedsiębiorstwo jest organizacją wielu oddziałów, posiada lub wykorzystuje łącza telekomunikacyjne i używa lub planuje budowę systemu wymiany informacji pomiędzy swoimi klientami za pomocą internetu, konieczny jest wybór odpowiedniego sche-

matu przepływu danych w systemie informatycznym. Na diagramach z rysunków 1,2,3 można zauważyć, że występują funkcjonalne bloki wykonujące te same lub podobne operacje, ale niestety są one odrębnymi elementami systemu informatycznego przedsiębiorstwa. Obsługa zdalna oddziałów i aplikacja *B2B* realizują funkcje złożenia zamówienia jako oddzielne moduły. Implementacja tej samej czynności złożenia zamówienia w kilku różnych aplikacjach jest rozwiązaniem nieekonomicznym. Utworzenie modułu warstwy pośredniej, z której korzystają wszystkie elementy systemu jest na pewno dużo lepszym rozwiązaniem. Warunkiem taniej i skutecznej implementacji jest utworzenie systemu trójwarstwowego.

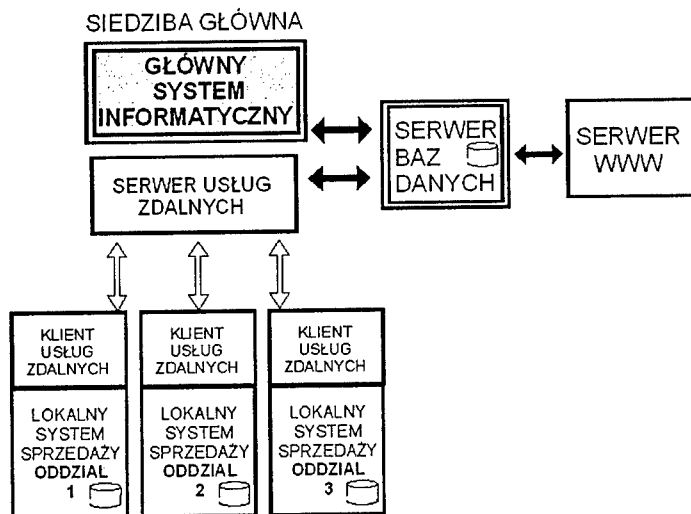
Dotychczas systemy informatyczne dla małych i średnich przedsiębiorstw były budowane przeważnie jako aplikacje dwuwarstwowe. W artykule został zaproponowany schemat zastosowania aplikacji trójwarstwowej jako podstawowego rozwiązania. Dużą przeszkodą w stosowaniu trójwarstwowej architektury systemu informatycznego były koszty odpowiednich modułów i ich duża komplikacja w powszechnym użyciu. Z chwilą przełamania tych barier powstaną proste narzędzia do budowy aplikacji trójwarstwowych i na pewno zostaną bardzo szybko wdrożone.

7. SYSTEM INFORMATYCZNY ZREALIZOWANY DLA FIRMY ŚREDNIEJ WIELKOŚCI

Do realizacji i rozbudowy systemu informatycznego średniej wielkości wielooddziałowej firmy dystrybucyjnej z branży IT (*Information Technologies*) zostały wykorzystane wytyczne podane w punkcie 3. System informatyczny, który składa się z kilku spójnych aplikacji został rozbudowany w pierwszej fazie o aplikację *B2B*. Aplikacja klasy *B2B* została uruchomiona na serwerze *IIS* i połączona z istniejącymi bazami danych systemu informatycznego działającego w przedsiębiorstwie. Wzory dokumentów zostały podane w programie Word i połączone poprzez XML z aplikacją *B2B*. Do pracujących programów została dodana nowa aplikacja zarządzająca użytkownikami systemu *B2B* i aplikacja, która konfiguruje układ oferty handlowej.

W dalszym etapie wielowarstwowa aplikacja obsługująca pracę oddziałów zamiejscowych została zamieniona na usługi terminalowe. Wyłączenie lokalnych systemów sprzedaży i obsługi klienta, zmniejszyło koszty eksploatacyjne systemu informatycznego. Kolejną fazą modyfikacji systemu informatycznego jest integracja czynności biznesowych wykonywanych w aplikacjach biurowych z całym systemem informatycznym przedsiębiorstwa.

Diagram przepływu danych przedstawia rys 4.



Rys. 4. Trójwarstwowy schemat wymiany danych w przedsiębiorstwie.

8. PODSUMOWANIE

Rozwój informatyczny zarządzania przedsiębiorstwem wymusza na producentach baz danych modyfikację swoich aplikacji w kierunku pełnej możliwości replikacji i implementacji systemów podwójnych transakcji. Obecnie pełne możliwości usług zapewniają aplikacje baz danych firm *Oracle Corp.* i *MS SQL* firmy *Microsoft*. Stwarzają one najbardziej bezpieczne i elastyczne możliwości pracy wielooddziałowej firmy przy wykorzystaniu łączy telekomunikacyjnych. Jednym z interesujących rozwiązań jest zastosowanie usług terminalowych. Pod względem kosztów i bezpieczeństwa pracy może to być bardzo atrakcyjne rozwiązanie. Można zastosować dowolną bazę danych i zaimplementować jeden system informatyczny dla całej firmy ze wszystkimi jej oddziałami. Wszyscy użytkownicy pracują na jednej bazie danych, nie ma potrzeby stosowania skomplikowanych i kosztownych usług sieciowych. Jedynymi kosztami dodatkowymi stają się wtedy licencje serwera terminalowego i zakup odpowiednich usług telekomunikacyjnych.

System informatyczny powszechnego użytku, musi składać się z wielu elementów. Powinien pozwalać na budowanie aplikacji wieloplatformowych, usług sieciowych, integrację serwerów aplikacji, serwerów internetowych i pocztowych. Powinien dawać możliwość przenoszenia logiki biznesowej do innych aplikacji projektowania systemów informatycznych i modelowania procesów zachodzących w przedsiębiorstwach. W wielu przypadkach powinien mieć architekturę trójwarstwową.

BIBLIOGRAFIA

- [1] Garlan D., Shaw M.: Introduction to software. School of Computer Science, Carnegie Mellon University, Pittsburgh 1994.
- [2] Jacobson I., Booch G., Rumbaugh J.: The Unified Software Development Process. Addison Wesley Professional 1999.
- [3] Bass L., Clements P., Kazaman R.: Software Architecture in Practice, 2/E. Addison Wesley Professional 2003.
- [4] Chmielarz W.: Systemy informatyczne wspomagające zarządzanie. Aspekt modelowy w budowie systemów. Agencja Wydawnicza Elipsa Warszawa 1996.
- [5] Kisielewski J., Sroka H.: Systemy informatyczne biznesu. Agencja Wydawnicza Placet Warszawa 1999.
- [6] Martin Folwer, Kendall Scott.: UML w kropelce. Oficyna Wydawnicza LTP Warszawa 2002.

**ARCHITECTURE OF INFORMATION SYSTEM FOR SMALL
OR MEDIUM ENTERPRISE****Summary**

Information system for small or medium firm must resolve many new problems in the electronic economy. Evolution of telecommunication involves growth possibilities of exchanging information between independent enterprises. Web Services, XML, WSDL, SOAP, DCOM permit for organizing fast exchange of data between various information systems. Many activities performed by people may be taken over by application. Information system has to be designed for operation on many hardware platforms. This paper discusses fundamentals of system's architecture developed for small or medium firms. In conclusion the trial has been undertaken to indicate new concepts and directions of programming environments being used for building information systems.

Beata Krawczyk-Bryłka, Michał Piotrowski

**Zakład Politologii i Socjologii, Politechnika Gdańska
Katedra Architektury Systemów Komputerowych, Politechnika Gdańska**

KOMPUTEROWE ARCHIWIZOWANIE PRZEBIEGU PRACY GRUPOWEJ*

Streszczenie

Przedstawiono metodę rejestracji i archiwizacji przebiegu przedsięwzięcia realizowanego poprzez grupę osób. Projektowanie systemu komputerowego oparto na modelu pracy zespołowej WPW (wkład – procesy – wynik), który dobrze opisuje czynniki warunkujące efektywność grupy oraz uwzględnia wieloaspektową obserwację uzyskiwanych rezultatów. W pracy zaprezentowano zarówno zakres rejestrowanych przez system danych jak i sposób ich archiwizowania.

1. WPROWADZENIE

Współczesne warunki realizacji przedsięwzięć różnego typu narzucają konieczność posługiwania się ogromną ilością informacji, często pochodzącej z odległych dziedzin i stale zmieniającej się. Dodatkowo presja czasu towarzysząca decyzjom o wprowadzaniu nowoczesnych, innowacyjnych rozwiązań, które mają szansę przebicia się na konkurencyjnym rynku, powoduje stały wzrost zainteresowania czynnikami warunkującymi pracę grupową. Istnienie grupy społecznej zakłada wzajemne oddziaływanie co najmniej dwóch jednostek, współzależnych od siebie i posiadających poczucie wspólnej tożsamości. Prowadzone na szeroką skalę możliwości wykorzystania grup do poprawienia efektywności realizowanych zadań pozwoliły wyszczególnić główne cechy, jakimi powinna charakteryzować się grupa, by tworzyć skuteczny zespół:

- posiadanie wspólnego celu;
- nawiązywanie relacji społecznych pomiędzy członkami grupy;
- dobrze zdefiniowane role poszczególnych uczestników;
- komplementarne umiejętności członków zespołu;
- zaangażowanie w realizację wspólnego celu ogólnego oraz celów szczegółowych;
- poczucie odpowiedzialności uczestników za realizowane zadanie [1,2].

* Pracę wykonano w ramach grantu KBN nr 6T11 2003C/06098 („Rozwój środowisk przetwarzania zespołowego zapewniający obsługę sytuacji wyjątkowych”)

Obserwację powyższych czynników ułatwia model WPW (wkład – procesy – wynik) [1], który określa związek elementów będących punktem wyjścia pracy zespołowej z zachodzącymi w jej trakcie procesami oraz ostateczną efektywnością wykonania zadania.

W pracy zaprezentowano model WPW oraz przedstawiono implementację systemu wspomagającego komputerową rejestrację przebiegu pracy grupowej. Z uwagi na złożoność realizacji procesów oraz ich różnorodny charakter, a także wielość aspektów pełna rejestracja faktów nie jest możliwa. Dlatego też głównym zadaniem systemu archiwizacji jest zapamiętanie faktów najistotniejszych, których analiza umożliwi odtworzenie relacji istniejących w pracy grupowej opisanych modelem WPW.

2. OPIS MODELU WPW

2.1. Wkład w zespół

Przebieg i rezultat pracy grupowej uzależniony jest od następujących czynników: typu zadania, składu zespołu, warunków funkcjonowania zespołu.

Aby przedsięwzięcie motywowało członków grupy do działania powinno być dla każdego z nich zadaniem ważnym z punktu widzenia celów osobistych, grupowych czy też organizacyjnych lub ogólnospołecznych. Podporządkowanie zadania jasno określönemu celowi ogólnemu, który stanowi o istocie stworzenia zespołu, powinno zakładać sformułowanie celów szczegółowych. Ich zadaniem jest wyznaczanie kolejnych kroków w dążeniu do celu ostatecznego, określenie dynamiki pracy grupy oraz uzyskanie różnorodności wykonywanych zadań. Cele szczegółowe pozwalają też kontrolować realizację poszczególnych etapów zadania, są więc podstawą do formułowania informacji zwrotnych, koniecznych do podtrzymywania zaangażowania uczestników grupy.

Idealny skład zespołu zakłada, że jego liczebność nie przekroczy 20 osób, jednakże optymalna jest liczba od 6 do 10 uczestników, gdyż zapewnia aktywność wszystkich. Dużo bardziej złożona jest decyzja o doborze osób do grupy, który powinien zakładać kompletność umiejętności należących do trzech kategorii:

- wiedza specjalistyczna,
- umiejętności rozwiązywania problemów i podejmowania decyzji – dostrzeganie wskaźników problemu i możliwości rozwiązania go, ocena różnych wariantów, wybór decyzji,
- umiejętności interpersonalne – przekazywanie informacji, rozwiązywanie konfliktów, konstruktywne krytykowanie, asertywność, aktywne słuchanie [2].

Szczególnie dwie ostatnie kategorie oraz role grupowe przyjmowane przez uczestników w trakcie realizacji przedsięwzięcia są ściśle związane z predyspozycjami osobowymi, które decydują o przebiegu procesów grupowych opisanych w punkcie 2.2.

Warunki funkcjonowania zespołu to np. kultura organizacji w jakiej dana grupa realizuje swoje zadanie (czyli styl zarządzania, sposoby premiowania efektywności, akceptacja struktury zespołowej) czy też charakterystyka otoczenia społeczno – biznesowego działania grupy (np. zmienność środowiska, wartości preferowane w kulturze).

2.2. Procesy w pracy zespołowej

Procesy towarzyszące realizacji zadania przez grupę można rozpatrywać z kilku punktów widzenia. Pierwsze podejście zakłada, że głównym procesem grupowym jest roz-

wój zespołu rozwiązującego problem. Proces ten obejmuje kilka faz, które dotyczą nowo tworzonych grup zadaniowych:

1. forming – kształtowanie się zespołu, rozpoznawanie zadania, wybór metod pracy, pozyskiwanie potrzebnych informacji;
2. storming – powstawanie wewnętrznych konfliktów, różnicy zdań, emocjonalny opór przed zadaniem wynikający z różnych punktów widzenia;
3. norming – rozwijanie współpracy poprzez akceptującą wymianę poglądów, ustalanie norm dyskusji i rozwiązywania problemów;
4. performing – otwartość relacji, elastyczność zachowań, odpowiedzialność za wykonanie zadania, poszukiwanie rozwiązań i wdrażanie ich [3].

Drugie podejście skupia się na procesach, bez których zespół nie może funkcjonować [1], podając za najważniejsze:

1. przywództwo – może opierać się na formalnym autorytecie osoby odgórnie przedstawianej w grupie jako lider lub też na samoistnie wyłaniających się liderach zależnie od zadania (podkreśla się skuteczność takich właśnie autonomicznych zespołów w warunkach stale zmieniającego się otoczenia);
2. komunikacja – wraz ze wzrostem liczby współpracujących osób, rośnie stopień kompleksowości relacji pomiędzy nimi, co powoduje coraz większą złożoność komunikacji (często wspomaga się ten proces narzędziami IT, by usprawnić gromadzenie, przetwarzanie i przechowywanie informacji niezbędnych do realizacji zadania);
3. podejmowanie decyzji – proces o zasadniczym znaczeniu dla jakości wykonywanych zespołowo zadań, obejmujący cztery główne etapy:
 - zdefiniowanie problemu;
 - określenie możliwych rozwiązań;
 - ewaluacja rozwiązań;
 - wdrażanie wybranego rozwiązania;

Podejmowanie decyzji w grupie jest narażone na wiele pułapek ograniczających racjonalność wyniku. Główną jest tzw. efekt myślenia grupowego, czyli preferowanie jednego (zwykle pierwszego) wariantu decyzyjnego oraz duża odporność na krytykę, nasilany przez spójność grupy, jej izolację czy też stres. Kolejna to efekt polaryzacji, oznaczający podejmowanie bardziej skrajnych decyzji niż początkowo proponowane przez indywidualnych uczestników (np. bardziej optymistyczne przewidywania finansowe). Nie zawsze więc potoczne stwierdzenie „Co dwie głowy to nie jedna” ma pozytywne znaczenie [4];

4. spójność i klimat w zespole – wzajemna sympatia członków grupy ułatwia przebieg pozostałych procesów i ma pozytywny związek z efektywnością działań. Elementami sprzyjającymi budowaniu spójności i korzystnego nastawienia kooperantów jest wspólna wizja, poczucie bezpieczeństwa, wzajemna pomoc, tworzenie symboli identyfikujących grupę oraz jej orientacja zadaniowa.

Istotne z punktu widzenia tworzenia systemu komputerowego jest rozpatrywanie procesu grupowego jako kolejnych etapów realizacji konkretnego zadania. I tak w przypadku, gdy celem zespołu jest osiągnięcie kompromisu na drodze negocjacji, kolejne etapy współpracy to [5]:

1. przygotowanie do negocjacji – obejmuje ustalenie celu, zakresu ustępstw, taktyki i spraw administracyjnych przebiegu rozmów (jest to etap realizowany zwykle niezależnie przez każdego z negocjatorów);

2. otwarcie rozmów – jej zadaniem jest budowanie przyjaznej atmosfery pomiędzy rozmówcami, ustalenie procedur prowadzenia rozmów oraz poznanie partnera negocjacji (jego decyzyjności, oczekiwań);
3. przedstawienie propozycji – negocjacje właściwe, dla których charakterystyczne jest warunkowe formułowanie własnych oczekiwań w zakresie poszczególnych zagadnień negocjacyjnych, udzielanie wzajemnych ustępstw i licytowanie się;
4. zakończenie negocjacji – etap starannego sprawdzenia co zostało ustalone, ostateczne formułowanie ustaleń.

2.3. Wynik pracy zespołu

Efektywność pracy zespołu można rozpatrywać z punktu widzenia racjonalności podjętych w procesie współpracy decyzji, w kontekście rozwoju organizacji którym dany zespół jest podporządkowany, czy wreszcie z szerszej perspektywy ogólnospołecznej, której podjęte i wdrożone decyzje mają służyć. W wypadku, gdy praca grupy przybiera formę negocjacji, jej skuteczność oceniać należy w co najmniej trzech aspektach [6]:

1. skutek osiągnięty ich prowadzeniem – obustronna korzyść wynikająca z uczciwości stron i gwarantująca kontynuację relacji;
2. wydajność, czyli sprawność przeprowadzenia rozmów, ich parametry czasowe;
3. stosunki między stronami biorącymi udział w rozmowach.

Wyniki uzyskiwane w czasie realizacji zadania bywają jednak oceniane na różnym poziomie przez indywidualnych uczestników zespołu. Dlatego też wydaje się konieczne, by ostateczną opinię o skuteczności współpracy opierać na parametrach wynikających z subiektywnych ocen indywidualnych, konfrontacji wyników z wyznaczonym celem oraz z pozycji osoby niezaangażowanej w pracę grupy, czyli niezależnego eksperta. Co więcej istotny jest sposób dochodzenia do ostatecznych wyników, w tym kolejność składanych propozycji i kontrpropozycji. Zapamiętany przebieg pracy grupowej powinien posiadać taką własność, że pewne jego aspekty mogłyby być ponownie odtworzone i ponownie analizowane.

3. MONITOROWANIE PRZEBIEGU PRZEDSIĘWZIĘCIA

Zgodnie z przyjętym modelem WPW przeprowadzane eksperymenty dotyczące prostych ludzkich przedsięwzięć jak negocjacje są dzielone na dobrze zdefiniowane etapy. Na koniec każdego etapu następuje jego zatwierdzenie i zapamiętanie zdarzeń, które w jego trakcie były rejestrowane. Zależnie od potrzeby rejestracje danego etapu mogą być szczegółowe (np. dokładny przebieg dyskusji) lub skupiać się na jego podsumowaniu (np. wyniki negocjacji).

Narzucenie formalnego przebiegu badań społecznych może znacząco wpływać na ich przebieg i dodatkowo zmniejszać elastyczność tworzenia i wykonywania nowych typów eksperymentów. Dużo wysiłku wymaga takie jego zaprojektowanie, aby pozwalało na elastyczną i zarazem spójną formę rejestracji różnorodnych badań.

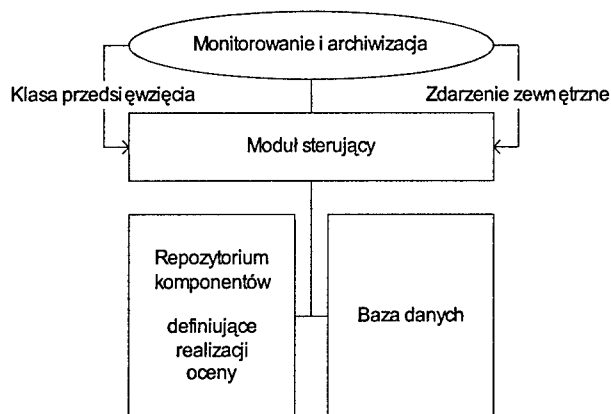
W dalszych rozważaniach ograniczymy się do dobrze zdefiniowanych trzech typów (klas) aktywności (przedsięwzięć):

- negocjacje typu kupno-sprzedaż;
- negocjacje rekrutacyjne;
- ocena i uszeregowanie pewnych obiektów (ranking).

Każda z tych klas może zawierać wiele różnych przypadków, różniących się np. celem negocjacji czy innym zestawem obiektów do uszeregowania. W naszym systemie zostało to rozwiązane w następujący sposób, że wszystkie badania definiujemy w postaci przedsięwzięć, które określają cel aktywności grupowej, ilość uczestników i przypisane im zadania.

Każde przedsięwzięcie jest przypisane do jednej z klas badań. Ze względu na to, że różne typy badań mogą się różnić przebiegiem, dla każdej klasy badań system umożliwia zdefiniowanie innego jego przebiegu. Dlatego też dla każdego przebiegu badania należy utworzyć odpowiednie komponenty i umieścić je w repozytorium komponentów (rys. 1) [8, 9]. Komponenty podzielone są na trzy kategorie:

- definiujące – używane do przygotowywania przedsięwzięć: opisów, tabel oceny wyników itp.;
- realizacji – używane w czasie rejestrowania przebiegu eksperymentu;
- oceny – używane do obliczania i prezentowania wyników eksperymentów.



Rys. 1. Komputerowa implementacja modelu WPW

Komponenty definiujące i oceny dołącza się do systemu dopisując je do odpowiedniego pliku konfiguracyjnego. Moduł sterujący dołączy je automatycznie w systemie do menu użytkownika.

Trochę inna jest sytuacja jeśli chodzi o komponenty realizacji. Tworzy się z nich odpowiednie zestawy, z których każdy reprezentuje jeden typ przebiegu i następnie definiuje się na bieżąco kolejność ich wykonywania. Taki zestaw może być składany z dowolnych komponentów dostępnych w repozytorium. Dzięki temu, że zestaw komponentów realizacji może zawierać odnośniki do innych zestawów, możliwe jest wykorzystanie wcześniej zdefiniowanych etapów badań i nie trzeba wielokrotnie pisać kodu, jeśli część etapów w różnych przebiegach przedsięwzięć jest taka sama lub podobna. Po zdefiniowaniu zestawu komponentów informację o nim dołączamy do bazy danych, gdzie znajduje się tabela z odpowiednim mapowaniem klasa przedsięwzięcia -> nazwa zestawu komponentów. Moduł sterujący korzystając z tej informacji uruchomi właściwy zestaw komponentów, zależnie od klasy przedsięwzięcia.

W zaprojektowanym systemie mamy aktualnie dwa sposoby rejestracji wyników aktywności grupowej (wyniki negocjacji i ranking obiektów). Mimo tego komponent z etapem rejestracji wyników jest wspólny. Może tak być dlatego, że zdefiniowaliśmy

klasę, po której dziedziczą wszystkie komponenty służące do rejestracji wyników eksperymentów [7].

Klasa ta ma dobrze zdefiniowany interfejs, pozwalający w jeden sposób wykorzystywać komponenty podrzędne. Dodatkowo ta klasa jest napisana wg wzorca projektowego factory (fabryka). Znaczący to, że obiektów (instancji) komponentów podrzędnych nie tworzymy bezpośrednio, tylko przy pomocy klasy nadrzędnej. Ona „wie” jakiego typu komponentu użyć w danym przypadku, a my nie musimy się tym już martwić. W celu zarejestrowania wyników aktywności grupowej musimy jedynie wywołać funkcję `showUserFillForm()`. Ona już zajmie się przeprowadzeniem całego etapu rejestracji wyników. Kiedy rejestracja będzie zakończona metoda zwraca wartość logiczną `prawda`. Wtedy komponent wywołujący wie, że może np. przejść do kroku następnego. Dzięki temu krok wypełniania wyników w naszym systemie może być taki sam dla wszystkich klas przebiegów badań.

4. REJESTROWANIE WYNIKÓW PRZEDSIĘWZIĘĆ

Sposób rejestracji wyników poszczególnych typów badań aktualnie obsługiwanych w systemie, zależy od klasy badań. Tworząc przedsięwzięcie danego typu definiujemy ogólne opisy ról, które potem czytają badani. W tych opisach są wymienione elementy podlegające negocjacji. Dodatkowo do każdej roli definiuje się tabele, które formalnie opisują różne warianty wyników negocjacji (rys. 2). Każdemu elementowi podlegającemu negocjacji przypisuje się zakresy wartości danego elementu (np. wysokości miesięcznego wynagrodzenia). Każdemu takiemu zakresowi przypisuje się wartość punktową. Jest ona tym wyższa im korzystniejsza dla danej roli jest wynegocjowana wartość. Tabele te na ogół nie są znane osobom negocjującym.

Rekr. kandydat (v2)					
Mies. wynagrodzenie NETTO		Data rozpoczęcia pracy		Wysokość corocznej premii	
2000-2500 zł	0	1 XI	0	do 2%	0
2500-3000 zł	40	15 X	60	2-4 %	100
3000-3500 zł	80	1 X	120	4-6 %	200
3500-4000 zł	120	15 IX	180	6-8 %	300
powyżej 4000 zł	160	1 IX	240	powyżej 8%	400
- Nie uwzględniono -	0	- Nie uwzględniono -	0	- Nie uwzględniono -	0
Miejsce wykonywania pracy		Długość kontraktu		Świadczenia socjalne	
Warszawa	0	0,5 roku	0	do 1000 zł	0
Gdańsk	80	5 miesięcy	30	1000-1500 zł	60
- Nie uwzględniono -	0	3 miesiące	60	1500-2000 zł	120
		2 miesiące	90	2000-3000 zł	180
		1 miesiąc	120	2000-3000 zł	240
		- Nie uwzględniono -	0	- Nie uwzględniono -	0

Ok

Rys. 2. Przykładowa tabela reprezentująca formalny opis różnych wariantów wyników negocjacji

Po zakończeniu etapu lub na koniec na koniec wykonania eksperymentu każdy uczestnik badania wypełnia formularz, w którym opisuje swój pośredni lub ostateczny wynik negocjacji. Mając tabelę punktową (z rys. 2) i zapisany wynik negocjacji możemy

wyznaczyć obiektywną skuteczność negocjacji. Tak wyznaczone skuteczności możemy porównywać na różne sposoby:

- porównanie wyników uzyskanych przez uczestników jednego eksperymentu;
- procentowe wyznaczenie skuteczności względem maksymalnej możliwej do uzyskania w danym badaniu dla każdego z uczestników;
- porównanie uzyskanej przez uczestników różnych eksperymentów skuteczności;
- porównanie skuteczności uczestników w zależności od wyznaczonych ich cech osobowych i in.

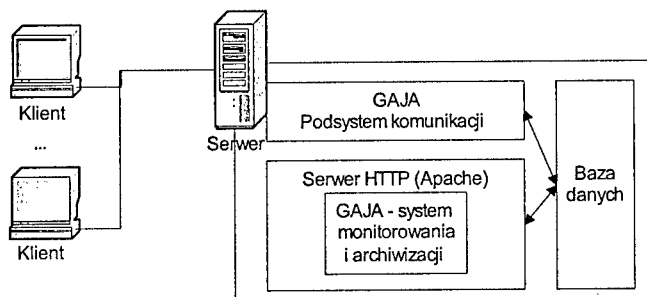
W przypadku badań związanych z rankingiem różnych obiektów (opinia społeczna) zapisujemy wyłącznie samo uszeregowanie obiektów. Tak zapisane uszeregowania możemy porównywać np. względem wyznaczonych cech osobowych bądź pod względem indywidualnego, bądź grupowego uszeregowania obiektów.

Poza opisanymi wyżej obiektywnymi wynikami badań rejestrujemy też subiektywne odczucia uczestników po odbytej dyskusji (negocjacji). Te subiektywne odczucia są rejestrowane przy pomocy ankiety zawierającej pytania typu: „Czy jesteś zadowolony z wyniku negocjacji?”. Na każde z pytań uczestnik udziela odpowiedzi w dziewięciostopniowej skali: od „Tak” przez „Raczej tak” itd. do „Nie”. Każdemu z pytań przypisana jest cecha, której to pytanie dotyczy, tzn. skuteczność, przyjazność, płynność, zaś każdej odpowiedzi jest przypisana odpowiednia wartość, np. odpowiedź „Tak” na powyższe pytanie ma wartość 9 a odpowiedź „Nie” ma wartość 1.

W ten sposób możemy wyznaczyć subiektywne cechy odbytych dyskusji w ramach badanej aktywności grupowej. Znając maksymalne możliwe do uzyskania wartości każdej cechy możemy wyznaczyć ich procentowy wynik w danym eksperymencie. Dodatkowo ankietę oceniającą przebieg przedsięwzięcia może wypełnić osoba nadzorująca eksperyment (ocena eksperta z zewnątrz). Daje to dodatkowy materiał, który można wykorzystać w dalszych analizach.

5. PODSUMOWANIE

Do opisu przedsięwzięć grupowych wykorzystano model WPW, który umożliwia wyróżnienie kilku istotnych etapów ich realizacji. Każdy z etapów może być komputerowo monitorowany i archiwizowany. W tym celu zbudowano system, którego architekturę przedstawia rys. 3.



Rys. 3. Architektura systemu Gaja

Jest to system internetowy oparty o architekturę klient-serwer. Klientem może być dowolna przeglądarka internetowa implementująca technologie JavaScript i Java Applet.

Serwer składa się z dwóch modułów. Główny moduł to aplikacja w PHP udostępniana przez serwer HTTP Apache. Drugi moduł to napisana w Javie aplikacja, zajmująca się przekazywaniem i rejestrowaniem komunikacji między uczestnikami badań o ile następuje ona przy pomocy komputera. Aktualnie obsługiwaną formą komunikacji w naszym systemie jest komunikacja tekstowa, czyli czat. Następnym krokiem będzie wykorzystanie możliwości komunikacji audiowizualnej.

System Gaja rejestruje przebieg badań w odpowiednio zaprojektowanej bazie danych, wyświetla wyniki badań i pozwala na przeprowadzenie ich analizy.

Podstawowe cechy takiego systemu przeznaczonego do archiwizowania i analizy danych pracy grupowej w badaniach społecznych to:

- bardzo dobra modyfikowalność systemu pozwalająca na łatwe wprowadzanie nowych typów badań, które system będzie mógł rejestrować;
- zastosowanie podejścia komponentowego przy tworzeniu systemu owocuje bardzo łatwą możliwością ponownego użycia kodu;
- oddzielone typy badań, co pozwala na niezależne rozwijanie i ewentualne poprawianie części systemu odpowiedzialnych za daną klasę eksperymentów. Wprowadza to też dodatkowe uporządkowanie w wewnętrznej architekturze systemu;
- szeroki zakres rejestrowanych danych pozwalający na przeprowadzanie bardzo różnorodnych analiz.

Umożliwi on zarejestrowanie wyników wielu eksperymentów. Ich analiza pozwoli na znaczącą rozbudowę modelu WPW dla wyjaśnienia znaczących aspektów pracy grupowej.

BIBLIOGRAFIA

- [1] Unsworth K.L., West M. A.: *Zespoły: wyzwania pracy w grupie* [w] Chmiel N. (red.): *Psychologia pracy i organizacji*. GWP, Gdańsk 2003
- [2] Katzenbach J.R., Smith D.K.: *Sila zespołów. Wpływ pracy zespołów na efektywność organizacji*. Oficyna Ekonomiczna, Kraków 2001
- [3] Kożusznik B.: *Zachowania człowieka w organizacji*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2002
- [4] Tyszką T., Zaleśkiewicz T.: *Racjonalność decyzji*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2001
- [5] Kamiński J.: *Negocjowanie. Techniki rozwiązywania konfliktów*. Poltext, Warszawa 2002
- [6] Nęcki Z.: *Negocjacje w biznesie*. Antykwa, Kraków 2000
- [7] Gellersen H., Gaedke M.: *Object-oriented Web Application Development*. W: IEEE Internet Computing January/February 1999
- [8] McInnis K.: *Internet, E-commerce and Component Architectures*. Castek 1999, <http://www.cbd-hq.com>
- [9] Gellersen H., Wicke R., Gaedke M.: *WebComposition: an object-oriented support system for Web engineering lifecycle*. W: Computer Networks and ISDN Systems 29, 1997

COMPUTER BASED ARCHIVIZATION OF GROUP ACTIVITIES PROCESSES

Summary

The method of recording and archiving of sequence of activities realised by work team are described. Computer system design was based on groupwork model: contribution – processes – outcome (CPO), accurately describing factors influencing effectiveness and taking into account multicriterial analysis of results. Recording and archiving capabilities of the system are presented.

Jerzy Nawrocki, Krzysztof Pospiech, Maciej Dorsz

Instytut Informatyki, Politechnika Poznańska

PROINFO: SYSTEM ZARZĄDZANIA INFORMACJĄ O PRZEDSIĘWZIĘCIU PROGRAMISTYCZNYM

Streszczenie

W trakcie trwania przedsięwzięcia programistycznego powstaje wiele różnych dokumentów, zwłaszcza jeśli pragnie się być w zgodzie ze standardami IEEE, modelem CMM/CMMI itp. Opracowanie szerokiego zestawu dokumentów jest czasochłonne i drogie. Ponadto, informacje zawarte w różnych dokumentach pokrywają się, co może prowadzić do niespójności całej dokumentacji. Ideą systemu ProInfo realizowanego w Instytucie Informatyki Politechniki Poznańskiej jest unikanie naruszeń spójności poprzez identyfikację "jednostek informacyjnych" (np. adresu e-mailowego analityka) występujących w całym zestawie dokumentów projektowych, określanie powiązań między dokumentami lub ich fragmentami (ang. traceability), przechowywanie wersji tych jednostek w bazie danych przedsięwzięcia i generowanie dokumentów na podstawie jednostek informacyjnych oraz szablonów (z każdym standardem IEEE związany byłby szablon).

1. WSTĘP

W przedsięwzięciach informatycznych informacja o produkcie, jak też i o samym przedsięwzięciu odgrywa istotną rolę. Informacja ta jest najczęściej gromadzona w postaci dokumentów. Liczące się na świecie organizacje, takie jak IEEE, ANSI, czy ISO, opracowały szereg standardów dotyczących dokumentowania przedsięwzięć i produktów. Różne dokumenty mają sporo elementów wspólnych.

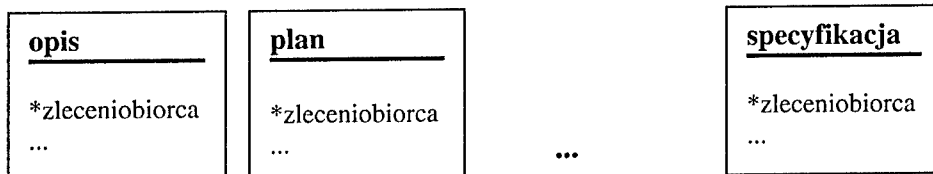
Powielanie informacji w różnych dokumentach wynika z chęci uczynienia ich kompletnymi (ang. *self-contained*) i czytelnymi. **Kompletność** oznacza, że dokument powinien zawierać wszelkie informacje dotyczące danego aspektu i nie powinien (nadmiernie) odwoływać się do innych dokumentów. **Czytelność** dokumentu ma wiele interpretacji; chodzi także o łatwość znalezienia szukanej informacji. Na przykład specyfikacja wymagań, zgodnie ze standardem IEEE 830 [1], musi zawierać krótki opis produktu na wysokim poziomie abstrakcji. Opis taki występuje również w Dokumencie Rozpoczęcia Przedsięwzięcia (ang. *Project Initiation Document*, PID) wykorzystywanym w metodzie PRINCE 2 [2].

Powielanie informacji ma oczywiście swoje wady. Ta sama praca (np. krótki opis produktu) jest wykonywana wielokrotnie zwłaszcza, gdy autorami różnych dokumentów są różne osoby (np. specyfikację wymagań opracowuje analityk a PID – kierownik przedsięwzięcia [5]). Jeszcze istotniejszym problemem jest utrzymanie spójności zestawu dokumentów, szczególnie w obliczu częstych zmian. W Programowaniu Ekstremalnym [3] problem ten rozwiązano w sposób ekstremalny: dokumentację ograniczono do kodu i przypadków testowych. Niestety, w wielu przedsięwzięciach takie rozwiązanie jest niedopuszczalne. Inne podejście polega na oddzieleniu informacji od dokumentów. Informację (dane o produkcie i przedsięwzięciu) umieszcza się w bazie danych, natomiast dokument jest raportem generowanym na żądanie na podstawie danych zawartych w bazie danych. Takie podejście zastosowano na przykład w systemie Rational RequisitePro służącym do zarządzania wymaganiami [4]. Niestety, RequisitePro jest ograniczony tylko do zbierania i dokumentowania wymagań.

System ProInfo budowany w Instytucie Informatyki Politechniki Poznańskiej ma służyć do zarządzania pełną informacją o przedsięwzięciu i ma obejmować nie tylko specyfikację wymagań, ale także zarządzanie przedsięwzięciem, opis architektury itp. W rozdziale drugim artykułu przedstawiono ideę sposobu oddzielenia jednostki informacyjnej od dokumentu, wprowadzanie danych oraz ukazano generowanie dokumentów. W kolejnej sekcji skupiono się na zarządzaniu zmianą. Omówiono korzyści płynące z wersjonowania jednostek informacyjnych oraz możliwości ich wzajemnego powiązania. Rozdział czwarty przedstawia przykładowe role występujące w ramach przedsięwzięcia programistycznego, jak również ukazuje sposób ich rozróżnienia w systemie ProInfo. Następna sekcja omawia inicjowanie systemu oraz przedstawia modyfikację już istniejących szablonów. W rozdziale szóstym pokrótce zaprezentowano architekturę omawianego narzędzia. Część siódma dotyczy porównania systemu ProInfo z funkcjonalnie podobnymi już istniejącymi rozwiązaniami. Rozdział ósmy prezentuje podsumowanie systemu oraz nakreśla dalsze kierunki prac nad ProInfo.

2. WPROWADZANIE DANYCH I GENEROWANIE DOKUMENTÓW

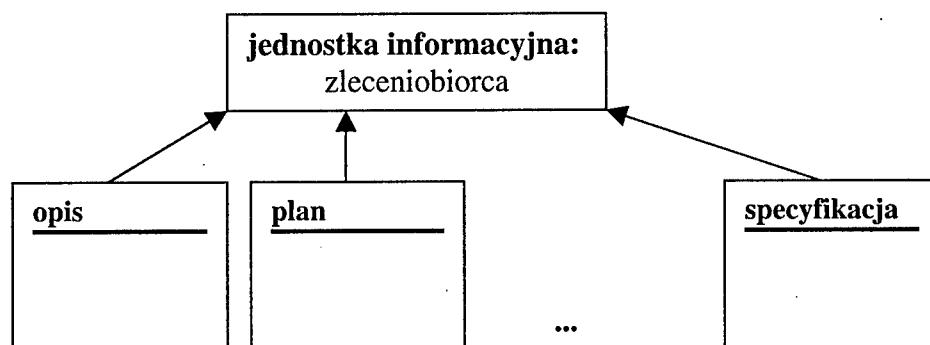
Zgodnie ze standardem IEEE 12207 [6] informacja o zleceniobiorcy (ang. *issuing organization*) powinna znaleźć się w dokumentach typu: opis, plan, procedura, zapis, raport, specyfikacja. Oznacza to duplikację informacji w wielu niepowiązanych ze sobą dokumentach. Podejście takie ilustruje rysunek 1.



Rys. 1. Duplikacja jednostki informacyjnej.

W przeciwieństwie do ww. modelu system ProInfo umożliwia współdzielenie jednostki informacyjnej przez szereg dokumentów. Czyni ją niejako niezależną od dokumentu. Podejście takie pozwala spojrzeć na dokumenty jako byty podrzędne w stosunku do jednostek informacyjnych. Dzięki temu, zamiast niezależnych dokumentów i w efekcie

niepowiązanych ze sobą danych, powstaje nowy model zależności między jednostkami informacyjnymi a dokumentem. Tworzy się hierarchia, w której elementami nadrzędnymi są jednostki informacyjne. Powiązania takie przedstawiono na rysunku 2.



Rys. 2. Oddzielenie jednostki informacyjnej od dokumentu.

Oddzielenie jednostki informacyjnej od dokumentu wyklucza duplikację danych oraz zapewnia utrzymanie spójności i aktualności dokumentów. W modelu, w którym dokumenty zależą od jednostek informacyjnych, generowanie dokumentacji na żądanie zapewnia jej aktualność i spójność. Ponadto wystarczy jednokrotne wprowadzenie danej jednostki informacyjnej do systemu, co oszczędza wiele niepotrzebnej i monotonnej pracy. Pozwala to uczestnikom przedsięwzięć programistycznych na skoncentrowanie się na dostarczaniu informacji do systemu i uwalnia ich od uciążliwości związanych z pisanem i ciągłym aktualizowaniem obszernej dokumentacji projektowej.

Informacje do systemu ProInfo można wprowadzać na dwa sposoby. Pierwszy polega na wypełnianiu formularza internetowego generowanego przez system. Drugi sposób polega na wczytaniu pliku XML z danymi. Format tego pliku jest zależny od rodzaju informacji. W niniejszym artykule omówimy, bardziej przyjazne dla użytkownika, podejście pierwsze.

Tworzenie szablonu, na podstawie którego generowane będą dokumenty, odbywa się przy pomocy narzędzi systemowych pozwalających określić, jakie jednostki informacyjne chcemy zawrzeć w danym dokumencie. System zapisuje dokumenty jako pliki XML. Daje to możliwość bardzo elastycznego sterowania wyglądem dokumentów, dzięki czemu można je dopasować do przyjętych wcześniej przez organizację standardów. Wygląd dokumentu jest opisany za pomocą arkusza stylów XSL. Można zatem w łatwy i szybki sposób zmienić wygląd takiego dokumentu.

Należy podkreślić, że system ProInfo abstrahuje od konkretnego standardu tworzenia dokumentacji. Szablony dokumentów operując na jednostkach informacyjnych, dają swobodę i możliwość generowania ich zgodnie z dowolnym standardem tworzenia dokumentacji. Dzięki temu organizacja programistyczna posługująca się np. normą ISO 9001:2000 lub CMM/CMMI, która zaprojektowała szablony dokumentów zgodnie z przyjętym modelem, może, np. na życzenie klienta, wygenerować dokumentację wg np. standardu IEEE 830. Wystarczy, aby przygotowała szablony dokumentów zgodnie ze specyfikacją IEEE 830. Co więcej, nie nastąpi duplikacja jednostek informacyjnych, co

oznacza, że generowana na żądanie dokumentacja wewnątrz firmy, jak również dla klienta, będzie zawsze spójna i aktualna.

3. ZARZĄDZANIE ZMIANĄ

Z punktu widzenia użytkownika zmiana powinna wyglądać następująco: właściciel jednostki informacyjnej (np. kierownik przedsięwzięcia) określa jednostkę mającą ulec zmianie, wprowadza zmianę i zmiana jest widoczna we wszystkich dokumentach zawierających tę jednostkę informacyjną. Z punktu widzenia systemu proces ten jest nieco bardziej rozbudowany i oferuje użytkownikowi dodatkowe możliwości. Jedną z nich jest możliwość określenia powiązań między danymi (*ang. traceability*). Wówczas modyfikacja jednostki informacyjnej pociąga za sobą konieczność zweryfikowania innych z nią powiązanych jednostek. Znacząco ułatwia to sposób śledzenia zależności w poszczególnych artefaktach.

Zmiana jednostki informacyjnej pociąga za sobą konieczność analizy spójności nowej wersji jednostki z pozostałymi jednostkami (na przykład zmienione wymaganie może stać się niespójne z przygotowanymi wcześniej testami lub implementacją). Istnieją dwie strategie analizy spójności: "z wyprzedzeniem" i "na żądanie". W analizie "z wyprzedzeniem" system natychmiast bada na jakie jednostki informacyjne ta zmiana będzie miała wpływ i prosi użytkownika o dokonanie ich aktualizacji. Przy analizie na żądanie sprawdzanie spójności jest odłożone do momentu generowania dokumentu. dopiero wtedy system bada powiązania i prosi o podjęcie akcji lub sugeruje wygenerowanie poprzedniej spójnej wersji danego dokumentu.

Jest też możliwość sprawdzenia spójności danych systemu na życzenie użytkownika. Wówczas wszystkie powiązania zostaną przeanalizowane pod kątem zaistniałych zmian, a użytkownik zostanie poinformowany o wyniku i będzie mógł podjąć stosowne akcje.

Wspomagany przez system proces zmiany ma również inne zalety. Zwalnia użytkownika z typowo biurowej pracy. Nie jest już konieczne wprowadzanie takich danych jak autor zmiany, data modyfikacji itp., gdyż te informacje system jest w stanie sam zapewnić. W ten sposób dodatkowo zabezpieczamy proces przed ludzkimi błędami.

Utrzymywanie dokumentacji zostało również uproszczone. Kontrolowany proces zmian pozwala na uzyskiwanie dokumentacji z określonego okresu. Wystarczy podać żadaną datę, a system sam wygeneruje dokumenty z odpowiednią dla tej daty zawartością. Znika problem pamiętania, gdzie i u kogo znajduje się jedna z poprzednich wersji dokumentu. Pozwala to skupić uwagę na innych istotnych aspektach pracy nad rozwojem produktu.

4. ROLA UŻYTKOWNIKÓW I ICH UPRAWNIENIA

System umożliwia w ramach jednego przedsięwzięcia dodawanie użytkowników oraz grup użytkowników. Role i grupy będą się różniły uprawnieniami w zakresie dostępu do informacji o przedsięwzięciu oraz wykonywania określonych czynności.

System jest projektem pilotowym i będzie wykorzystywany w ramach Studia Rozwoju Oprogramowania [5] na Politechnice Poznańskiej.

Najważniejsza rola przypada Kierownikowi Projektu [5], gdyż to on jest odpowiedzialny za dane przechowywane w systemie. Jego rolę można opisać jako Administratora projektu. W zakres jego odpowiedzialności będzie wchodziło nadawanie praw programistom oraz innym użytkownikom do potrzebnych im informacji oraz funkcji.

Dodatkowo będzie istniała jeszcze jedna rola, w której wystąpi Opiekun SDS [5]. Jest on koordynatorem projektów wykonywanych w ramach Studia Rozwoju Oprogramowania. Będzie miał dostęp do wszystkich aktywnych projektów w celu oceny ich zaawansowania. Ta rola nie zakłada jakiegokolwiek modyfikowania danych w ramach przedsięwzięć, raczej tylko odczyt w postaci dokumentów przygotowywanych przez Kierowników Projektów. Będzie miał jednak uprawnienia do stworzenia nowych projektów w systemie oraz dopisania do nich użytkowników administrujących w ramach danego projektu.

5. INICJOWANIE I MODYFIKACJE SYSTEMU PROINFO

Najwygodniejszy sposób inicjacji systemu ProInfo polega na utworzeniu szablonej struktury jednostek informacyjnych oraz powiązanych z nimi raportów.

Ponieważ ProInfo umożliwia współdzielenie jednostek informacyjnych, każdą z nich wystarczy wprowadzić tylko jeden raz. Rozpatrzmy prosty przykład. Użytkownik zdefiniował dwa szablony: A i B, z których każdy zawiera identyczną sekcję opisującą wszystkie osoby uczestniczące w przedsięwzięciu. Następnie po wypełnieniu pierwszego dokumentu (A), użytkownik rozpocznie wprowadzanie jednostek informacyjnych poprzez kolejny dokument (B). Nie będzie jednak musiał wykonywać żmudnej pracy wprowadzania dublujących się danych powtórnie, gdyż po otwarciu drugiego dokumentu (B), sekcje te wypełnią się automatycznie.

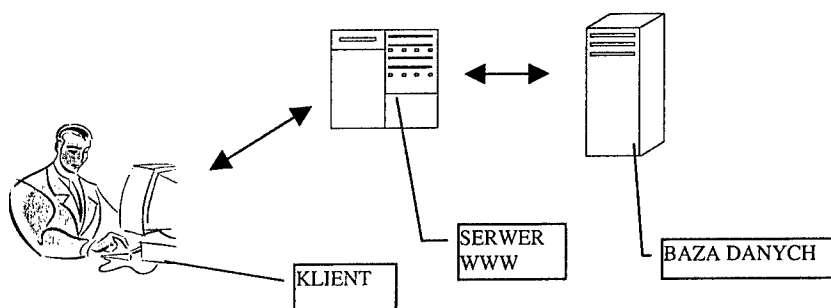
Możliwe jest również wprowadzanie informacji do systemu bez uprzedniego przygotowania wyrafinowanych szablonów dokumentów. Wystarczy wówczas utworzyć uproszczony formularz zawierający niezbędne sekcje. Warto jednak jeszcze raz podkreślić, że raz wprowadzone informacje można wykorzystać wielokrotnie. Oznacza to, że późniejsze opracowanie szablonów w oparciu o już zdefiniowane i wprowadzone jednostki informacyjne, spowoduje ich automatyczne wypełnienie w wygenerowanym dokumencie.

System zezwala na modyfikację już istniejących jednostek informacyjnych. Umożliwia też określenie daty ważności jednostek informacyjnych. Dzięki temu dokument wygenerowany wstecz, będzie miał zawsze jednakowy wygląd. Wersjonowanie danych zapewni również jego odtwarzalność w okresie trwania przedsięwzięcia. Jest to niezwykle przydatne np. w czasie audytu ISO, podczas którego audytorzy mogą poprosić o kompletną dokumentację projektową na dany dzień.

6. ARCHITEKTURA SYSTEMU PROINFO

System ProInfo jest typu Open-Source i powstał w oparciu o popularne rozwiązania stosowane do rozwoju aplikacji internetowych. Baza danych jest oparta o MySQL, jako serwer WWW działa APACHE a język programowania to PHP.

Architektura systemu jest trójwarstwowa: serwer WWW, baza danych oraz klient uzyskujący dostęp poprzez przeglądarkę internetową.



Rys. 1. Architektura systemu ProInfo

Instalacja systemu ogranicza się do zainstalowania wspomnianych aplikacji oraz stworzenia bazy danych z wymaganymi relacjami początkowymi. Dodatkowo system umożliwi dodanie do nowo tworzonych projektów schematów danych oraz dokumentów przygotowanych już wcześniej, np. pod kątem zastosowania w ramach SDS [5].

Otwartość na modyfikacje powinna wpłynąć na zwiększenie funkcjonalności systemu oraz dać możliwość usprawniania jego działania w celu spełniania potrzeb przyszłych użytkowników.

7. PROINFO A INNE ROZWIĄZANIA

Implementację systemu ProInfo poprzedzono analizą porównawczą istniejących narzędzi podobnego przeznaczenia..

W trakcie analizy wykorzystano następujące kryteria:

- Szablony: Czy dany system pozwala na utworzenie szablonu dokumentu, a następnie na wielokrotne jego wykorzystanie do generowania dokumentów (zwanych również raportami),
- Współdzielenie danych: Czy dany system pozwala na współdzielenie poszczególnych jednostek informacyjnych przez różne szablony,
- Zależności: Czy dany system zezwala na określenie i śledzenie zależności między jednostkami informacyjnymi, tak by w razie potrzeby poinformować użytkownika, że być może korzysta z nieaktualnych informacji,
- Wersjonowanie: Czy dany system wersjonuje jednostki informacyjne i umożliwia wygenerowanie np. poprzedniej wersji dokumentu,
- Sposób wprowadzania danych: W jaki sposób następuje wprowadzanie danych do systemu.

Wyniki analizy porównawczej, przedstawione w tablicy 1, pokazują, że – przynajmniej z punktu widzenia przyjętych kryteriów – budowa systemu ProInfo ma sens, gdyż nie ma systemu, który spełniałby postawione cele. Użycie np. języka skryptowego do osiągnięcia zamierzonej funkcjonalności oznacza równocześnie niemiernie większy nakład pracy, komplikację procesu tworzenia szablonów oraz zmianę profilu użytkowników systemu na osoby dobrze zaznajomione z informatyką. Dodatkowym atutem ProInfo jest sposób wprowadzania danych – 1 bezpośrednio poprzez szablon lub nawet poprzez wygenerowany dokument. Taką cechą charakteryzują się jedynie narzędzia firmy DGA, przy czym nie umożliwiają one np. współdzielenia danych.

Tablica 7.1

Zestawienie narzędzi umożliwiających zarządzanie dokumentami

Kryterium	ProInfo	Rational SODA	Business Object	Crystal Reports	Narzędzia firmy DGA	ID Solutions
Szablony	√	√	√	√	√	√
Współdzielenie danych	√	√	√	√	Nie	√
Zależności	√	*	**	nie	***	*
Wersjonowanie	√	*	**	√	√	√
Sposób wprowadzania danych	Poprzez formularz	Np. RequisitePRO, Rational Rose	pliki tekstowe; Excel; bazy danych, itp.	pliki tekstowe; Excel; bazy danych, itp.	Poprzez szablony	Bazy danych: SQL, XML ...

Legenda:

* - śledzenie zależności między danymi można uzyskać poprzez porównanie dat modyfikacji danych przy pomocy języka skryptowego; analogicznie można wygenerować dokument dla danych archiwalnych, przy czym należy wówczas odpowiednio zbudować cały szablon dokumentu

** - jak w *, przy pomocy VisualBasic

*** - jak w *, przy pomocy np. JavaScript

8. ZAKOŃCZENIE

W niniejszej pracy przedstawiono pilotowy projekt systemu ProInfo wspomagającego zarządzanie dokumentacją poprzez wydzielenie jednostek informacyjnych. System ten wspomaga zarządzanie wiedzą o przedsięwzięciu oraz uwalnia użytkowników od poświęcania sporych ilości czasu na prace biurowe dotyczące tworzenia dokumentacji. System umożliwia generowanie dokumentacji na podstawie jednostek informacyjnych oraz sprawdza powiązania między nimi. Projekt zostanie wdrożony w ramach Studia Rozwoju Oprogramowania PP w przyszłym roku. Do najbliższych prac rozwojowych należy będzie poszerzenie możliwości systemu o zarządzanie różnymi rodzajami danych – nie tylko tekstowych.

Praca została wykonana w ramach grantu BW-91-399/04.

BIBLIOGRAFIA

- [1] IEEE Std 830-1998, Software Engineering Standards Committee of the IEEE Computer Society
- [2] CCTA, Managing Successful Projects with PRINCE 2, The Stationary Office, London, 2002.
- [3] Programowanie Ekstremalne, <http://www.extremeprogramming.org/>
- [4] IBM Rational RequisitePro, <http://www-306.ibm.com/software/awdtools/reqpro/>
- [5] Nawrocki J., Pawałowski P., Pospiech K., System Internetowy Wspomagający Zarządzanie Przedsięwzięciami Metodą Prince 2, Zeszyty Naukowe Politechniki Gdańskiej 2003
- [6] IEEE Std 12207-1995, Industry Implementation of International Standard ISO/IEC 12207 : 1995

PROINFO: INFORMATION SYSTEM IN SOFTWARE ENTERPRISE**Summary**

Many different documents are created during programming projects, especially if they are to follow IEEE standards, CMM/CMMI model, etc. Construction of wide range of documents is time consuming and expensive. Moreover, information gathered in different documents may be duplicated, which can cause inconsistency of whole documentation. The idea of ProInfo system built at Institute of Computing Science at Poznan University of Technology is to avoid inconsistency through identification of "information units" (such as analyst's e-mail address) used in all project documents and construction of links between documents or their elements (traceability). All occurring versions of information units are stored in project's database. Basing on these units and document templates (each IEEE standard would have a corresponding template) documents are generated.

Katarzyna Pardo, Wojciech E. Kozłowski

INFO TECH

**ZMIANY OTOCZENIA BIZNESOWEGO A NAKŁAD PRACY
W PROJEKCIE INFORMATYCZYM – STUDIA PRZYPADKÓW**

Streszczenie

Dobra i ścisła współpraca z klientem jest jednym z najważniejszych warunków powodzenia przedsięwzięcia informatycznego. Z tego powodu zmienne wymagania klienta powinny być w projekcie odzwierciedlane możliwie szybko i bez negatywnego wpływu na organizację wewnętrzną samego projektu. Studia przypadków przedstawione w artykule wskazują typowe zjawiska występujące w projektach prowadzonych w ścisłej współpracy z klientem. Dzięki wyróżnieniu i uogólnieniu typowych zjawisk możliwe staje się planowanie przedsięwzięć z założeniem i uwzględnieniem zmienności wymagań klienta zarówno w samym planie projektu, jak i w nastawieniu zespołu projektowego do wykonywanych zadań.

1. WSTĘP

Projekt informatyczny, aby osiągnął sukces, nie może być prowadzony w oderwaniu od środowiska zewnętrznego [1]. Na środowisko to składają się zarówno organizacje będące klientami firmy prowadzącej projekt jak i sytuacja na rynku krajowym i globalnym, czy też zmiany w polityce i prawodawstwie.

W firmie realizującej projekty w charakterze podwykonawcy w systemie *outsourcingu* wszystkie służące realizacji zadań prace mogą i muszą być prowadzone w ścisłym kontakcie z klientem, ponieważ odpowiedź na jego zmieniające się potrzeby musi być szybka i konkurencyjna w stosunku do innych istniejących na rynku firm konsultingowych.

1.1. Sytuacja małej firmy w zmiennym otoczeniu biznesowym

INFO TECH jest małą gdańską firmą informatyczną, angażującą się w ambitne projekty w dziedzinie systemów wbudowanych (ang. *embedded systems*) dla dużych zagranicznych kontrahentów [2]. Unikalność firmy wynikająca z wiedzy i doświadczenia zespołu jak i znajomości procesów marketingowych, projektowych i produkcyjnych jej głównych kontrahentów pozwala klientom znajdować w niej wykonawcę niestandardowych przedsięwzięć. Klienci eksperymentujący z nowymi technologiami we współpracy z INFO TECH poszukiwać mogą nowatorskich i skuteczniejszych rozwiązań.

Dzięki temu firma może rozwijać swoje *know-how* wraz z każdym nowym projektem. Problemy do rozwiązania charakteryzują się bowiem często dużą złożonością i spekulatywnym charakterem. Projekty mają charakter **systemowy**, nierzadko wymagający **integracji rozwiązań z różnych dziedzin**. Zespół firmy z konieczności szybko absorbuje wiedzę na temat nowych technologii i dziedzin ich zastosowań, aby móc być dostawcą konkurencyjnych rozwiązań.

Mamy tutaj jednak do czynienia również z drugą stroną medalu. **Eksperymentalny charakter niektórych projektów** powoduje, że wiedza na temat wykorzystywanej technologii może być niepełna w momencie rozpoczęcia pracy. Prowadzi to do **częstych zmian wymagań** w rezultacie weryfikacji uzyskanych przez klienta rozwiązań cząstkowych i prototypowych. Skutkiem tego jest praca wykonywana często pod naciskiem terminów dla kolejnych etapów weryfikacji rozwiązania oraz trudności ze spełnieniem wszystkich wymagań jakościowych oczekiwanego produktu.

Bezpośredni i częsty kontakt z klientem umożliwia szczegółową wymianę informacji i pozwala klientowi „trzymać rękę na pulsie”, ale tym samym klient ma możliwość dużej ingerencji w to, co dzieje się po stronie podwykonawcy [3]. Współpraca w typowych przypadkach zakłada **integrację procesów tworzenia produktu w firmach klienta i podwykonawcy** poprzez ulokowanie zarządzania biznesowego docelowym produktem w firmie klienta, powierzenia zarządzania projektem tandemu osób delegowanych przez obie strony oraz możliwości samodzielnego wyboru optymalnego modelu realizacji projektu przez podwykonawcę. Ten **3-warstwowy system** zakłada niewątpliwie, że przebieg projektów może się zmieniać w zależności od podejmowanych w ustalonych terminach decyzji biznesowych klienta, wynikających zarówno z oceny stanu realizacji projektu jak i zmieniającej się sytuacji rynkowej oraz finansowej. Elementem wymagającym szczególnej uwagi jest w tej sytuacji konieczność **dynamicznego przydzielania pracownikom nowych zadań** w trakcie trwania projektu, czego skutkiem może być nierównomierność rozkładu zadań w momentach kluczowych, jak np. przy zbliżającym się terminie weryfikacji rozwiązania. Ponadto możemy mieć do czynienia z tym, że pracownicy tracą z oczu ogólny cel całego projektu, koncentrując się jedynie na powierzonych im zadaniach, nie będąc przy tym pewnymi, jak bardzo wynik ich pracy będzie przydatny w całości. Może się więc pojawić problem z motywacją zespołu.

W zmieniającym się tak dynamicznie środowisku trudno jest przedstawić jednoznaczny przepis na sukces. A jednak firma może pochwalić się nieprzerwanym wzrostem na przestrzeni ostatnich pięciu lat. Wynika to zarówno ze stabilnej i rozwijającej się współpracy z tradycyjnymi klientami poprzez uwzględnienie INFO TECH w planach kolejnych przedsięwzięć jak i uzyskania niezbędnych w tym zakresie działalności referencji, pozwalających na zdobycie zaufania nowych klientów. Wzrost liczby realizowanych równolegle prac pozwala na rozwój przedsiębiorstwa, a więc systematyczne powiększanie zespołu. Ponadto przeprowadzone w zeszłym roku badania wskazują, że czas poświęcany w projektach na poszczególne fazy jest w przybliżeniu stały i zgodny z modelem **COCOMO** [4], co stanowić może potwierdzenie obecności w firmie dobrych praktyk menedżerskich.

Jak więc przygotować się do szybkiej i skutecznej odpowiedzi na zmieniające się potrzeby klientów, a przy tym nie zdeformować organizacji procesu wytwórczego? Jak utrzymać kompetencję i dobre praktyki nie tylko w tak zmiennym środowisku zewnętrznym, ale również w okresie rozwoju firmy wymuszając właściwą rekrutację i wdrażanie nowych pracowników do zespołu oraz budowanie własnej bazy wiedzy?

2. WPROWADZENIE DO STUDIÓW PRZYPADKÓW

Niniejszy artykuł przedstawia dwa studia przypadków spośród kilku badanych w ostatnim czasie w firmie. Studia te mają na celu wskazanie w projektach informatycznych typowych zjawisk, które są skutkiem ścisłej współpracy z klientem, a więc przystosowywania się na bieżąco do jego zmiennych wymagań.

2.1. Zbieranie danych – system raportowania pracy

W roku 2002 w ramach pracy dyplomowej został w firmie wykonany i wdrożony system raportowania pracy w projektach. System ten pozwala użytkownikom (menedżerom, kierownikom projektów oraz członkom projektów) łatwiej planować projekty informatyczne, szczegółowo raportować pracę wykonaną w ramach tychże projektów oraz generować sumaryczne raporty wewnętrzne i zewnętrzne w formacie MS Word.

Dane w postaci przepracowanych przez zespół godzin zebrane dzięki systemowi raportowania służą zarówno do bieżącego nadzorowania stanu prowadzonych projektów jak i analizy *post-mortem* projektów już zakończonych. Jako studia przypadków do niniejszego artykułu wybrano dwa projekty, na temat których dane zgromadzone zostały w systemie raportowania.

2.2. Klasyfikacja – typowe czynności w projektach

Aby rozpocząć studia przypadków, należy podać pewne założenia dotyczące nazewnictwa i oznaczeń. Wzorując się na modelu RUP [5] wprowadzono pojęcie **czynności** (ang. *activity*), którym określać będzie się np. analizę wymagań, projektowanie czy też implementację.

Projekty prowadzone w firmie INFO TECH różnią się od siebie zestawem wykonywanych w nich czynności. Wynika to z założeń przyjętych dla projektów – wiele z nich posiada narzucony przez klienta zestaw nazw czynności, według których powinna być raportowana wykonywana praca. Niezależnie jednak od nazw poszczególnych czynności, analizując projekty zauważono, że można czynności te zakwalifikować do grup o nazwach powszechnie znanych w terminologii dotyczącej projektów informatycznych.

Zunifikowane nazwy czynności, które stosowane będą w opisie poszczególnych projektów znajdują się w poniższej tabeli. W opisie każdej z czynności podano, jakie działania będą do niej zaliczane.

Tablica 2.1

Podział czynności w projektach informatycznych

Nazwa czynności	Działania wchodzące w skład czynności
Requirements	Czynność ta zawiera w sobie wszystkie prace wykonywane przy definiowaniu zakresu projektu, a więc zarówno w fazie wstępnej (Pre-Study), jak i właściwego procesu zbierania wymagań (Requirements), a także wykonywanie różnego rodzaju specyfikacji, np. koncepcji produktu, rozwiązania, przeglądu technologii i rozwiązań dostępnych na rynku, specyfikacji środowiska pracy docelowego produktu jak i specyfikacji testu akceptacyjnego. Do czynności tej zaliczono również przygotowanie środowisk programistycznych i sprzętowych (Environment), głównie ze względu na poznawczy charakter tej czynności w przypadku wykorzystywania nowych technologii.

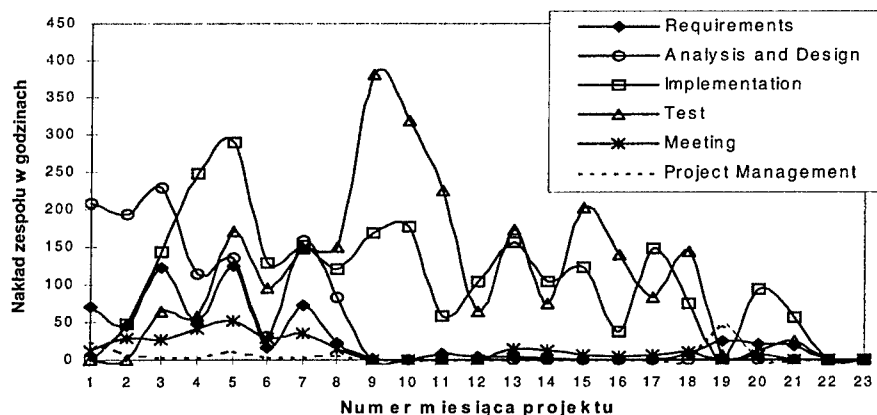
Analysis and Design	Do czynności tej zaliczono Analysis and Design oraz Business Modeling. Czynność ta obejmuje zarówno tworzenie modelu i architektury rozwiązania, jak i przygotowanie specyfikacji systemowych i funkcjonalnych opracowywanego produktu oraz odpowiednich do nich specyfikacji testów.
Implementation	Kategorię tę należy rozumieć jako tradycyjną implementację, w sensie przygotowania szczegółowych specyfikacji na poziomie modułowym oraz tworzenia (i dokumentacji) kodu. Czasem do tej czynności wchodzi również testy modułowe, wykonywane podczas implementacji.
Test	Test zawiera w sobie wszystkie czynności związane z testowaniem, zarówno w środowisku wykonawcy, jak i docelowym. Zaliczono tutaj również prace wdrożeniowe (Deployment) oraz czynności wykonywane przy istniejącym sprzęcie (Hardware).
Meeting	Spotkania w gronie pracowników firmy oraz spotkania z Klientem (inspekcje, prezentacje wyników, ocena stanu projektu).
Project Management	Do tej czynności zaliczone zostały te nie wymienione wcześniej, a dotyczące samego prowadzenia projektu: zarządzanie projektem (Project Management) i zarządzanie konfiguracją i zmianami (Configuration and Change Management).

3. STUDIA PRZYPADKÓW

Projekt Alpha jest dużym projektem software'owo-hardware'owym, w którym budowane jest kompletne urządzenie wraz z oprogramowaniem. Zaliczyć go można do kategorii projektów **eksperymentalnych**, w których klient poszukuje optymalnego rozwiązania poprzez wypróbowanie nowej technologii.

Projekt Beta jest w całości projektem software'owym, w którym również budowane jest od podstaw oprogramowanie dla opracowanego przez klienta urządzenia.

3.1. Projekt Alpha



Rys.1. Przebieg projektu Alpha

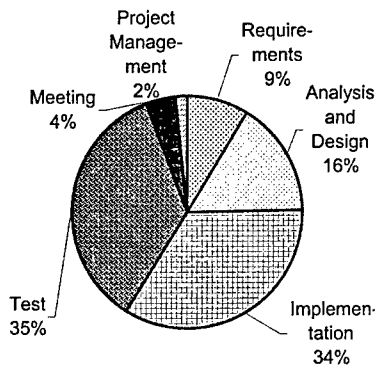
Ze względu na zakres oraz rozmiar projekt Alfa podzielony został na sześć osobnych części, z czego dwie pierwsze części – wytworzenie sprzętu i oprogramowania – prowadzone były równolegle. Początek projektu (pierwsze dwa miesiące) charakteryzuje się dużym skokiem czynności związanych z analizą i projektowaniem przy stosunkowo

niewielkim udziale wydobywania i specyfikacji wymagań. Dzieje się tak dlatego, że przed rozpoczęciem projektu Alfa przeprowadzono duży projekt typu *Pre-Study* (ok. 10% całości projektu Alfa) i to w nim skupiły się działania związane z wymaganiami jak i wyborem technologii.

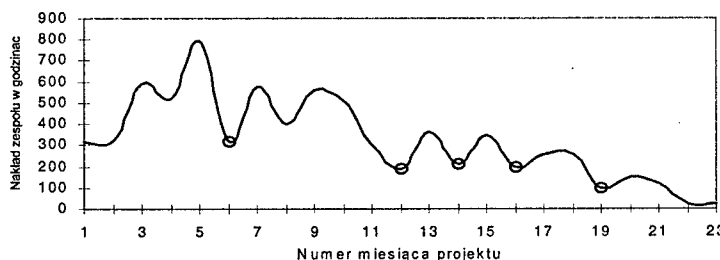
Trzeci miesiąc projektu przynosi weryfikację wymagań i wstępnego modelu rozwiązania po spotkaniach inspekcyjnych z klientem. Konieczne są rewizje modelu i stąd pojawienie się w tym okresie prac związanych z modelowaniem. W konsekwencji podjętych decyzji rusza na dobre implementacja wraz z testowaniem modułowym. Pierwsza wersja sprzętu i oprogramowania wykonana zostaje w okolicach 6. miesiąca projektu. Czynności następujące potem to weryfikacja wymagań i rewizja modeli. Miesiąc 8. to okres świąteczny, stąd widoczny jest przejściowy spadek nakładu pracy. Po przerwie świątecznej rozpoczyna się faza testów funkcjonalnych i systemowych (do 11. miesiąca projektu) wraz z implementacją wymaganych zmian.

Kolejne miesiące to testowanie prototypowego rozwiązania w laboratorium klienta oraz instalacja pilotowa prototypowego rozwiązania. W ślad za tym następuje budowa i uruchomienie małej serii urządzeń z uwzględnieniem nowych wymagań klienta jak i wniosków z instalacji pilotowej prototypu. W tym okresie daje się zauważyć wyraźne przeplatanie się czynności implementacyjnych z testowymi – to kolejne etapy eksperymentu i testowanie wytworzonych rozwiązań. W okolicach 19. tygodnia rozpoczyna się okres spotkań i formułowania wymagań dla kolejnego (wdrożeńowego) projektu oraz

tworzenie dokumentacji użytkownika. Projekt *Pre-Study* przeprowadzony przed projektem Alfa w celu zbadania możliwości technologicznych spowodował, że jedynie 25% całości głównego projektu poświęcono na specyfikację wymagań, analizę i projektowanie (Rys.2). Pozostała część projektu to w mniej więcej równym stopniu implementacja i testowanie kolejnych rozwiązań. Szczególnie widoczne jest to od 11. miesiąca projektu, kiedy prace koncentrowały się na testach, korektach i optymalizacji rozwiązania poprzez przygotowanie małej serii urządzeń na bazie wersji prototypowej. Przy tworzeniu tej małej serii wystąpiły jeszcze 3 iteracje, dzięki którym końcowe rozwiązanie odpowiadało już zmodyfikowanym oczekiwaniom klienta.



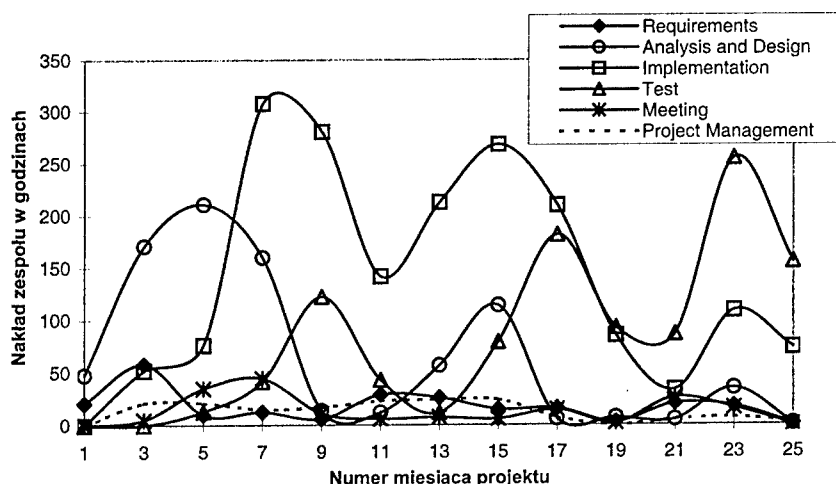
Rys.2. Czynności w projekcie Alpha.



Rys.3. Sumaryczny przebieg projektu Alpha.

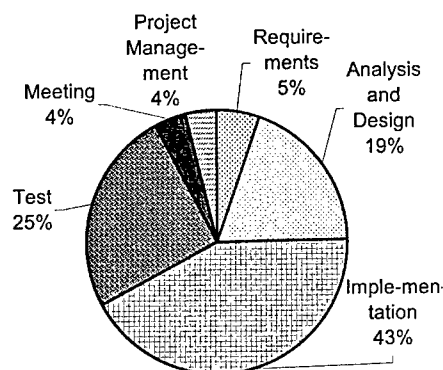
Rys. 3 przedstawia sumę nakładu na wszystkie czynności w projekcie przez cały czas jego trwania. Daje się zauważyć (miesiąc 2 i 8), że przerwy takie jak świąteczna czy urlopową wyraźnie odbijają się na sumie nakładu pracy w projekcie – spadek natężenia prac jest porównywalny z tym, jaki obserwujemy podczas przechodzenia między iteracjami (oznaczone na wykresie znakiem O).

3.2. Projekt Beta



Rys. 4. Przebieg projektu Beta

Projekt Beta został rozpoczęty w roku 2002 i trwa nadal. Na rys. 4., obrazującym przebieg projektu w czasie, wyraźnie wskazać można kolejne iteracje. Pierwsza z nich, zawierająca w sobie pełny cykl od zbierania i specyfikacji wymagań, poprzez analizę, projektowanie, implementację do testów, kończy się ok. 11. miesiąca projektu przygotowaniem pierwszej prototypowej wersji produktu. W tym okresie zaobserwować można powrót do specyfikacji wymagań oraz analizy i projektowania. Wynikał on z faktu, że klient stwierdził pojawienie się dodatkowych wymagań rynkowych i postanowił dokonać funkcjonalnej rozbudowy produktu. Nakład na czynności implementacyjne spada, aczkolwiek nie do zera. Faza testów następuje szybciej niż w poprzedniej iteracji, bo już w trzecim miesiącu od jej rozpoczęcia, a nie – jak poprzednio – ok. 7. miesiąca. Widoczne jest dużo mniejsze natężenie czynności związanych ze specyfikacją i projektowaniem w porównaniu do iteracji poprzedniej. Około 20. miesiąca projektu iteracja druga zbliża się do końca. Następuje kolejna konfrontacja rezultatów projektu z oczekiwaniami klienta, a w rezultacie tego decyzja o wdrożeniu przygotowanego rozwiązania do produkcji. Równocześnie klient podejmuje decyzję o opracowaniu zmodyfikowanej wersji urządzenia, którą mógłby zaoferować w pokrewnej dziedzinie zastosowań. W projekcie następuje zatem ponowne nasilenie działań związanych z wymaganiami oraz analizą i projektowaniem dla iteracji trzeciej, lecz w znacznie mniejszej skali niż w dwóch poprzednich iteracjach. Zaobserwowany w tym okresie wzrost działań związanych z testowaniem to budowa zautomatyzowanego środowiska testowego dla opracowanego urządzenia.



Rys. 5. Czynności w projekcie Beta

Automatyzacja procesu testowania jest częścią polityki firmy w celu podniesienia zarówno jakości jak i wydajności pracy w projektach o iteracyjnym charakterze.

Udział procentowy czynności w projekcie Beta jest podobny do tego, który występował w projekcie Alfa. Podobnie jak tam, czas projektu poświęcony na specyfikację wymagań i analizę to ok. 25% – ale w tym przypadku nie było projektu *Pre-Study*. Nakład pracy na implementację jest w związku z tym o ok. 10% większy niż w projekcie Beta, i o prawie 20% przewyższający dotychczasowy nakład pracy na testowanie (należy przypomnieć że projekt jeszcze się nie zakończył, zaś faza testów trzeciej iteracji właśnie trwa).

3.3. Zaobserwowane zjawiska

Wspólną cechą przedstawionych powyżej projektów jest ich iteracyjny charakter. Podstawowe zjawiska, które możemy wyróżnić w tego typu projektach to:

Zmniejszający się nakład pracy w kolejnych iteracjach. Widoczny szczególnie w projekcie Alfa, gdy po opracowaniu inicjalnego rozwiązania następuje jego dalszy rozwój oraz wytwarzanie kolejnych wersji produktu.

Płynna granica pomiędzy iteracjami. Nakład pracy pomiędzy iteracjami nie spada do zera. Iteracje nachodzą na siebie – ustalanie wymagań dla iteracji następnej nakłada się z ostatecznymi testami i oceną iteracji poprzedniej.

Niewielki udział testowania pod koniec pierwszej iteracji. Widoczny wyraźnie w obydwu studiach przypadków nieduży udział testów w pierwszej iteracji wskazuje na to, że iteracja pierwsza kończy się zamknięciem pewnego etapu implementacji i testów modułowych. Celem podstawowym jest uzyskanie prototypowego rozwiązania, które może być poddane pierwszej ocenie klienta. Pełne przetestowanie systemu nie jest zwykle ani możliwe, ani wskazane (technicznie i ekonomicznie) dopóki nie będzie wersji odpowiadającej zwykle zmodyfikowanym po ocenie prototypu oczekiwaniom klienta.

Zmniejsza się udział analizy i projektowania w kolejnych iteracjach. Zmiany wymagań rozpoczynające nową iterację nie mają już dużego wpływu na ogólny model systemu. Stąd też czynności w późniejszych iteracjach to głównie implementacja i testowanie – wynika to z faktu, że prace koncentrują się głównie na korektach i optymalizacji rozwiązania. Niewątpliwie natomiast jest znaczące powtarzanie się nakładu pracy na testowanie, stąd też niezwykle celowym wydaje się być włożenie wysiłku w zautomatyzowanie procesu testowania.

Duże znaczenie dla projektów mają zdarzenia dnia codziennego. Ze względu na charakter małej firmy zarówno okresy wolne od pracy jak i choroby członków zespołu mają duże znaczenie w skali całego projektu. Podobny efekt będą miały czasowe migracje członków zespołu do innych projektów.

6. PODSUMOWANIE

Iteracyjny charakter projektów wskazuje na pewne działania, które są konieczne przy planowaniu tego typu przedsięwzięć. Należy pamiętać o tym, że iteracje są nieuniknione, jeżeli poszukujemy rozwiązania najlepiej spełniającego oczekiwania klienta.

Planowanie powinno być więc wieloetapowe, ale niekoniecznie z naciskiem na dostarczenie końcowej wersji produktu w jak najkrótszym czasie. Ważne jest raczej położenie nacisku na możliwie szybkie ukończenie pierwszej iteracji, nawet bez pełnego przetestowania wytworzonego produktu. Dzięki temu w miarę wcześnie w projekcie zweryfikowana zostanie technologia oraz zdobyta wiedza konieczna do dalszego rozwoju produktu. Możliwe będzie również szybkie podjęcie przez klienta kolejnej decyzji biznesowej co do kierunku kontynuacji prac. Planowanie takie uwzględniać też powinno możliwość zupełnego zarzucenia wybranego rozwiązania i rozpoczęcia prac od nowa. Zespół projektowy powinien być więc przygotowany na to, że praca włożona w projekt może być – pozornie – niewykorzystana. Bardzo ważna jest więc dobra wymiana informacji pomiędzy wszystkimi udziałowcami projektu.

Projekty w firmie INFO TECH planowane są z reguły według Rational Unified Process. Daje się zauważyć, że mimo wszystko projekty planowane w ten sposób z czasem przybierają w poszczególnych iteracjach charakter wodospadu, gdzie wyraźnie zdefiniowane są terminy zakończenia poszczególnych etapów oraz produkty do dostarczenia. Planowanie etapowe natomiast sprawia, że projekty w naturalny sposób przyjmują charakter ewolucyjno-przyrostowy. W tego rodzaju elastycznych i ewolucyjnych projektach kluczowym czynnikiem jest ścisła współpraca z klientem. Dzięki temu po zakończeniu każdej iteracji może nastąpić wspólna i bezkonfliktowa, a więc konstruktywna ocena wykonanych prac oraz określone kierunki dalszego rozwoju produktów.

BIBLIOGRAFIA

- [1] Górski J. et al.: *Inżynieria oprogramowania w projekcie informatycznym*, Wydawnictwo „MIKOM”, Warszawa 2000.
- [2] Paulk, Mark C.: *Using the Software CMM® in Small Organizations*, Carnegie Mellon University, 1998.
- [3] Boehm, B., Bose, P., Horowitz, E., Lee, M., *Software Requirements Negotiation and Renegotiation Aids: A Theory-W Based Spiral Approach*, Seattle: IEEE Computer Society Press, 1995
- [4] Boehm, B.: *Software Engineering Economics*, Prentice Hall, 1981
- [5] Rational Unified Process: <http://www.rational.com/products/rup/index.jsp>

BUSINESS ENVIRONMENT CHANGES VS. WORK EFFORT IN SOFTWARE PROJECTS – CASE STUDIES

Summary

Successful cooperation with the customer is one of the main success factors in outsourced software projects. Thus the changing requirements of the customer should be taken into consideration as quickly as possible and without negative interference with the internal project organization. The article presents two case studies, where typical situations in close customer-supplier cooperation projects are described. Generalizing these typical situations allows to plan a software project in a better way and prepare the project team with respect to the ever-changing environment.

Justyna Stasieńko

Katedra Informatyki Gospodarczej, Szkoła Główna Handlowa

PROFIL KLIENTA W PROCESIE ZARZĄDZANIA RELACJAMI Z KLIENTAMI

Streszczenie

Tworzenie profili polega na podziale grup klientów na segmenty ludzi, którzy mają wspólne upodobania i zwyczaje zakupowe. Informacje demograficzne i informacje dotyczące zachowań wykorzystywane są do stworzenia pomocnego portretu klienta. W pracy zaprezentowano profil klienta w systemie zarządzania relacjami z klientami. Zaproponowany profil klienta uwzględnia informacje demograficzne, osobowość klienta, preferencje klienta oraz wartość klienta dla firmy. Personalizacja niesie ze sobą duże możliwości. Dla klientów oznacza łatwiejszy i szybszy dostęp do informacji, sprawniejszą obsługę, indywidualne dopasowanie ofert itp. Firmy posiadające profile swoich klientów stają się przyjaźniejsze, dobierają produkty, czy też usługi w sposób spersonalizowany dla klientów, znając preferencje tworzą odpowiednie programy marketingowe i lojalnościowe.

1. WSTĘP

Zarządzanie relacjami z klientami jest zjawiskiem stosunkowo nowym, nie mającym pełnej i jednoznacznej definicji, stąd też można się spotkać z przypisywaniem do systemów tej klasy zarówno rozbudowanych baz danych o klientach i rynkach, jak i zaawansowanych narzędzi informatycznych do wspierania działań marketingowych, akcji promocyjnych czy też programów lojalnościowych. W artykule tym zarządzanie relacjami z klientami traktowane będzie jako systemy wspomagające oraz odpowiednie narzędzia informatyczne, które umożliwiają ich skuteczne wdrożenie.

Klient pragnie odczuwać zainteresowanie ze strony firmy, z którą współpracuje. Oczekuje z jej strony inicjatywy, fachowości, partnerstwa, uczciwości w interesach oraz przestrzegania umów. Utrzymanie klienta staje się obecnie najważniejszym zadaniem w zarządzaniu relacjami z klientami. Znajomość klienta i jego potrzeb oznacza m.in. umiejętność przewidywania jego decyzji. Na decyzję klienta mają wpływ czynniki społeczne i psychologiczne. Do czynników społecznych zaliczane są: warstwa społeczna klienta, wartości kulturowe, grupa odniesienia, która stanowi wzorzec stylu życia, liderzy opinii, faza cyklu życia rodziny klienta itp.[1], zaś w skład czynników psychologicznych wchodzi m.in.: osobowość, motywacje, opinie, postawy, tendencja do ryzyka, innowacyjność.

Rosnąca konkurencja, zaspokajanie potrzeb klientów, zrozumienie klienta i jego potrzeb oraz zachowań, analizowanie statusu, lokalizacji itp. sprawia, że personalizacja, czyli uwzględnianie indywidualnych cech i potrzeb, które charakteryzują klienta, jest bardzo ważnym procesem. Najczęściej polega ona na zdefiniowaniu modelu klienta, który zawierałby szeroką wiedzę na temat klienta. Model jest także nazywany profilem, który reprezentuje preferencje i zainteresowania klienta, charakterystykę osobistą itp. [2].

W różnych źródłach klient jest inaczej postrzegany. K. Burnett [1] dzieli klientów na cztery kategorie: klienci lojalni, klienci konkurencyjni, klienci do przekonania oraz klienci lojalni wobec konkurencji. Klient może być indywidualną jednostką lub podmiotem gospodarczym. Klientem może być również użytkownik systemu. W referacie rozpatrywany jest indywidualny klient, jego preferencje, cele i osobowość.

W artykule omówiono zarządzanie relacjami z klientami, proces personalizacji oraz zaproponowano profil klienta indywidualnego, przydatnego w zarządzaniu relacjami z klientami, nie tylko w tradycyjnym rynku, ale i elektronicznym. Profil klienta posłuży klasyfikacji klienta do odpowiedniej grupy według tzw. „drabiny zażyłości”.

2. ZARZĄDZANIE RELACJAMI Z KLIENTAMI

Celem zarządzania relacjami z klientami jest maksymalizacja wartości przypadającej na klienta, która polega na zwiększaniu długofalowej wartości generowanej przez klienta oraz zmniejszaniu kosztów obsługi i komunikacji. Warunkiem koniecznym zaspokajania potrzeb klientów jest znajomość ich preferencji. Najważniejsze w stosunkach klient - firma jest zadowolenie klienta. Spełnianie oczekiwań klienta polega na: poznaniu oraz zrozumieniu jego potrzeb, wykorzystaniu tej wiedzy, poprawieniu komunikacji, zaoferowaniu klientowi tego, czego inni nie oferują, a także budowaniu więzi pomiędzy firmą a klientem.

Poznanie potrzeb klienta to ciągły proces zbierania danych z wielu źródeł. Dzięki tym danym wytworzony zostaje obraz kompleksowej wiedzy o kliencie. Na podstawie tej wiedzy tworzony jest profil klienta. Konieczne jest zachowanie integralności profilu klienta (poprzez ciągłe uzupełnianie i synchronizację danych). Aby zrozumieć jego potrzeby należy je przeanalizować. W ujęciu masowym odbywa się to m.in. za pomocą narzędzi business intelligence¹, segmentacji, wzorców zachowań, zaś w ujęciu spersonalizowanym wykorzystywana jest analiza profilu oraz obserwacja najbardziej wartościowych klientów. Należy „słuchać” klienta poprzez kontakty, ankiety itp. Należy dostrzec i analizować werbalizowane preferencje klienta. Wszystkie dopasowane dane o kliencie są podstawą do uruchamiania procesów personalizacyjnych, a tym samym akcji mających na celu zacieśnienie związku klienta z firmą. Polepszanie jakości zarządzania relacjami z klientem następuje wraz ze zwiększaniem się liczby istotnych danych w profilu, szczególnie deklarowanych przez użytkownika oraz wynikających z historii transakcji.

Klienci, którzy mają możliwość dotarcia do szybkiego i głębokiego wsparcia informacyjnego, czyli otoczeni spersonalizowaną informacją i usługami udostępnianymi w kontekście ich potrzeb, są bardziej zaangażowani. Przeprowadzane kampanie marketin-

¹ Business intelligence zwykle stanowi rezultat dogłębnych analiz szczegółowych danych firmy. Obejmuje technologie baz danych i aplikacji, a także techniki analityczne. Czasem używane jest jako synonim "wspomagania decyzji", chociaż business intelligence jest znacznie bardziej ogólne z technicznego punktu widzenia. Może obejmować między innymi zarządzanie wiedzą, ERP i zbieranie danych.

gowe i komunikacja z klientem powinny być prowadzone w sposób spójny, dopasowując się do preferencji klienta oraz opierając się na wiedzy o kliencie zawartej w profilu.

Klienci oczekują nowych rozwiązań, a nie tylko nowych produktów. Oferty powinny wykorzystywać wiedzę o kliencie i być zindywidualizowane. Aplikacje obsługowe i e-commerce mogą dostarczyć pełną wartość klientowi po otoczeniu ich warstwą aplikacji elektronicznego zarządzania relacjami z klientami.

Firma zyskuje w oczach klienta, który staje się jej lojalnym klientem, jeśli prowadzi proaktywne kontakty i doradztwo oparte na wiedzy o kliencie. Dzięki możliwości reakcji organizacji w czasie rzeczywistym w elektronicznym zarządzaniu relacjami z klientami, klient otrzymuje maksymalną możliwą szybkość obsługi. Spójność spersonalizowanej komunikacji we wszystkich kanałach elektronicznych (WWW, WAP, e-mail) dodatkowo zwiększa komfort klienta. Klient musi mieć poczucie bezpieczeństwa i panowania nad relacją. System musi pozwalać na natychmiastową modyfikację sposobu obsługi klienta.

3. PERSONALIZACJA

Tradycyjny marketing opiera się przede wszystkim na statystycznej segmentacji klientów, którzy grupowani są według np. wieku, wykształcenia, zarobków, stylu życia itp. Wybrany docelowy segment klientów jest wówczas anonimowy. I działania skierowane do wybranej grupy są anonimowe. Reklama, promocje, e-maile zawierają identyczną treść, bo nie są znane indywidualne preferencje klienta. Tradycyjny marketing okazuje się nie wystarczający, aby realizować cele przedsiębiorstwa. Globalizacja rynków i wzmożona konkurencja sprawiły, że przedsiębiorstwa zaczęły dostrzegać wartość indywidualnego klienta. Dbłość o klienta prowadzi do poznawania jego preferencji i potrzeb, podniesienia poziomu jego zadowolenia z oferowanych dóbr i usług, stworzenia silnej i emocjonalnej więzi z nim. Rodzi się rywalizacja o lojalność klienta. W obecnej gospodarce przedsiębiorstwa wiedzą, że sam produkt już nie wystarczy, aby klient pozostał stałym (lojalnym) klientem. Firmy muszą stosować nowe, coraz bardziej atrakcyjne środki marketingowe. W warunkach rosnącej konkurencji pojedynczy konsument staje się dla firmy wartością, gdyż o sukcesie zaczyna decydować indywidualne podejście do klientów, budowa stałej więzi z klientem i skuteczny proces wzajemnej komunikacji. Klient stał się więc najcenniejszym zasobem, o który zaczęto się troszczyć. Tak zrodził się proces personalizacji.

Personalizacja jest metodą zorientowaną na potrzeby klienta, uwzględnianie jego zachowania i preferencje, ukierunkowanych skłonności nabywczych, czy też zdefiniowanie osobistych wymagań. Personalizacja powinna stanowić podstawę każdego systemu zarządzania relacjami z klientami. Podstawowe działania związane z personalizacją to: udoskonalona lojalność klienta, udoskonalone zaufanie i satysfakcja klienta, identyfikacja użytkownika, udoskonalona użyteczność, wydajność i precyzja informacji, sprawność i informacyjna wydajność procesów, źródło dla „wycelowanych” promocji, indywidualny marketing.

Najważniejszą cechą personalizacji jest usatysfakcjonowanie klienta dostarczając mu właściwe treści, produkty i usługi. Personalizacja uwzględnia decyzje klienta i akceptuje jego preferencje. Powoduje, że nie jest on kolejnym adresatem masowego marketingu, lecz indywidualnym odbiorcą zorganizowanych działań. Personalizacja jest metodą zorientowaną na potrzeby klienta, co sprawia, że czuje się on kimś wyjątkowym. Personalizacja jest jednym ze sposobów walki o klienta.

Proces personalizacji jest wykorzystywany na coraz szerszą skalę w Internecie [[8],[9],[10]]. Firmy muszą wychodzić naprzeciw klientom, dlatego Internet i w biznesie

postrzegany jest jako rynek. Zarządzanie relacjami z klientami odbywa się też na płaszczyźnie elektronicznej.

Biorąc pod uwagę elektroniczny rynek, po zidentyfikowaniu klienta zostaje zaprezentowana mu spersonalizowana informacja (np. strona internetowa). To zachęca anonimowych użytkowników, aby zidentyfikowali się możliwie szybko w procesie np. zakupu. Z punktu widzenia dostawcy, ważne jest zachęcenie użytkownika do identyfikacji, aby móc mu dostarczyć możliwie dużą wartość dodaną wynikającą z personalizacji. Równie ważne jest przedstawienie wartościowej treści dla anonimowych użytkowników. Anonimowi użytkownicy nie mogą jednak otrzymywać spersonalizowanej zawartości, np. odpowiednio sprofilowanych promocji. Korzyści jakie odnosi klient z personalizacji to: możliwość składania zamówień na nietypowe produkty (indywidualne), korzystanie z promocji, sprawna i miła obsługa, szybki dostęp do poszukiwanych zasobów.

4. PROFIL KLIENTA

Precyzyjna analiza cech klienta tworzy jego indywidualny profil. Proces ten określany jest jako masowa indywidualizacja. Tworzenie profili polega na podziale grup klientów na segmenty ludzi, którzy mają wspólne upodobania i zwyczaje zakupowe. Profile są tzw. „kartami klienta”. Dla każdego klienta gromadzone są indywidualne dane. Obejmują one historie obsługi klienta, historie transakcji, wypełniane przez klientów różnego rodzaju ankiety, dane demograficzne, dane od partnerów, zachowania on-line. Dane klienta zawierają osobistą charakterystykę klienta. Mogą być one dostarczane przez samego klienta, jak i zebrane w sposób automatyczny, np. wywnioskowane na podstawie obserwacji. Dane klienta podzielić można na następujące kategorie: dane demograficzne, zainteresowania, preferencje klienta, cele i plany klienta.

Istotnym problemem w tworzeniu profilu klienta jest reprezentowanie wiedzy o zainteresowaniach, potrzebach klienta i zwyczajach zakupowe. Tradycyjnie, przedstawia się te zainteresowania w postaci zbioru słów kluczowych lub jako wektor słów z wagami. Pojawiają się także propozycje stosowania bardziej złożonych struktur opisujących wiedzę o preferencjach klienta [2]. Należą do nich: zbiór stereotypów, czyli opisów prototypowego użytkownika pewnej klasy użytkowników, charakteryzujących się wspólnymi zainteresowaniami oraz sieci semantyczne, które wskazują przedmiot zainteresowania z wyróżnieniem głównego przedmiotu zainteresowań oraz tematów z nim związanych.

Wyróżnia się kilka podejść wyznaczania profilu [3]. Jednym z nich są metody oparte o rejestrowanie wprost wyrażonego przez klienta przedmiotu zainteresowania w pewnym formularzu, bądź w postaci odpowiedzi na stereotypowe pytania. Innym podejściem jest analiza pytań kierowanych przez klienta do systemu (lub pracownika obsługującego danego klienta), uwzględniając częstość pojawienia terminu w pytaniu.

Profilem nazywamy obiekt p należący do zbioru P , gdzie P to zbiór wszystkich możliwych profili klientów. Profil p składa się z informacji demograficznych klienta², informacji na temat preferencji, osobowości oraz wartości klienta dla firmy.

$$p = \langle Id, Os, Pr, V \rangle \quad (5.1)$$

gdzie Id – informacje demograficzne,
 Os – osobowość klienta,
 Pr – preferencje klienta,
 V – wartość klienta dla firmy.

² Informacje demograficzne w innych źródłach rozbijane są na informacje demograficzne i psychograficzne [7]. Tu natomiast potraktowano je łącznie.

Każdy profil klienta (5.1) jest profilem indywidualnego klienta odnoszącego się do konkretnych zasobów informacyjnych (informacje demograficzne, informacje dotyczące osobowości, preferencji klienta oraz lojalności względem firmy). Każdy atrybut profilu p jest zbiorem wartości (mierzalnych i niemierzalnych).

$$Id = [id_1, id_2, \dots, id_n] \quad (5.2)$$

gdzie id_i oznacza każdą daną demograficzną, w skład których wchodzi takie dane jak: imię, nazwisko, wiek, płeć, zawód, wykształcenie, adres, PESEL, stan cywilny, status społeczny, dochód, styl życia, hobby, jeśli klient posiada rodzinę, to dochodzi liczba członków rodziny.

$$Os = [os_1, os_2, \dots, os_m] \quad (5.3)$$

Osobowość klienta O , przedstawiona jest jako wektor składający się z następujących m atrybutów: zwyczaje, przekonania, wierzenia, uczucia, inteligencja, zdolności, temperament, nawyki, motywy, zasady etyczne, uwarunkowania kulturowe, zmiany kulturowe.

Do preferencji klienta P , zaliczane są jego potrzeby i pragnienia.

Wartość klienta dla firmy V jest miarą lojalności klienta względem firmy w skali od 1 do 5. Wartość V pomoże w różnicowaniu klientów, od klientów kluczowych (najbardziej lojalnych i przynoszących dochody) do tych najmniej lojalnych.

Na podstawie modelu klienta w (5.1) można powiedzieć, że profil p stanowi zindywidualizowany zbiór informacji, bazując na indywidualnych informacjach demograficznych, osobowości, preferencji i wartości klienta dla firmy. Profil ten jest profilem dynamicznym w aspekcie zmieniających się danych. Grupowanie danych w odpowiednich bazach umożliwia szybkie sprawdzanie odpowiednich danych i proponowanie klientowi szczególnie dla niego przygotowanej oferty. Profil klienta pozwala zaklasyfikować klienta do odpowiedniej grupy. Do profilu p zastosowana została tzw. „drabina zażyłości” [11], która główny nacisk kładzie na V wartość klienta. Klient grupowany jest względem zażyłości klienta z firmą. Stopniami tej zażyłości są etapy: potencjalny klient (typ A), nabywca (typ B), klient (typ C), sojusznik (typ D) oraz adwokat (typ E). Obsługujący klienta pracownik, w zależności od stopnia zażyłości, czy inaczej mówiąc – lojalności, potrafi w sposób bardziej zorganizowany i profesjonalny obsłużyć klienta. Zaproponować najodpowiedniejszy produkt lub usługę. Podział taki pozwala także konstruować produkty lub usługi specjalnie „pod danego klienta”.

5. ZAKOŃCZENIE

„Głębszy” poziom personalizacji oparty jest na zrozumieniu indywidualnych personalnych cech klienta. Jest to nowa fala w polityce zarządzania relacjami z klientami. Korzyści z tego procesu są znaczne. Po pierwsze następuje wzrost satysfakcji i lojalności klientów. Po drugie następuje znaczący sukces w up-sellingu i cross-sellingu. Personalizacja obniża koszty operacyjne Call Center. Następuje również wzrost efektów kampanii i promocji. Koncepcja personalizacji leży w dużym stopniu na sumowaniu różnych danych o kliencie, tak aby stworzyć holistyczny obraz klienta. Konstruowanie profili zapewnia podstawę do zapoczątkowania tego, co handlowcy nazywają „dialogiem” z klientem.

W dalszej pracy nad profilowaniem klienta wprowadzony zostanie do systemu inteligentny agent. Inteligentny agent jest to mały, sprytny program pozwalający zautomatyzować wybraną czynność, często podejmujący decyzje w trakcie działania, reagujący w określonej sytuacji. Wprowadzenie do systemu zarządzania relacjami z klientami takiego

agenta zautomatyzuje pracę i usprawni kontakt z klientem. Na podstawie zebranych danych o klientach, ze sprzężenia zwrotnego domniemanego (niejawnego) lub jawnego, agent tworzy profile klientów i w sposób ciągły aktualizuje je.

BIBLIOGRAFIA

- [1] Burnett K.: *Relacje z kluczowymi klientami*. Oficyna Ekonomiczna, Kraków, 2002.
- [2] Indyka-Piasecka A.: *Propozycja personalizacji zapytań w internetowym systemie wyszukiwania informacji*, Materiały Konferencyjne MiSSI, 2002.
- [3] Bergtold A.: *Obsługa klienta – sprzedaż partnerska*.
<http://www.marketing.info.pl/biblioteka/artykuly/5152.htm>.
- [4] Newell F.: *Lojalność.com. Zarządzanie relacjami z klientami w nowej erze marketingu internetowego*. IFC PRESS Sp. z o.o., Kraków, 2002.
- [5] Conway K.: *Customer Connections. A Breakthrough Model for Analyzing and Influencing Customer Behavior*.
http://accountancyage.bitpipe.com/data/detail?id=1071510946_531&type=RES&x=1769912411.
- [6] Adamczyk J.: *Zbawienna personalizacja- czyli jak zdobyć lojalność e- klienta*.
<http://www.e-marketing.pl/artyk/artyk53.php>.
- [7] Tworzenie profili klientów i modeli. http://directmarketong.com.pl/html/artykuly/4/4_10.html.
- [8] Nguyen H.C.: *Preferencje użytkownika w systemach informacyjnych*, Materiały Konferencyjne MiSSI, 2000.
- [9] Sobecki J.: *Profile użytkowników w systemach informacyjnych dostępnych w sieci Internet.*, Materiały Konferencyjne MiSSI, 2002.
- [10] Fink J., Kobsa A.: *A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web*. User Modeling and User-Adapted Interaction 10(3-4), Special Issue on Deployed User Modeling, 209-249. <http://www.ics.uci.edu/~kobsa/papers/2000-UMUAI-kobsa.pdf>
- [11] Chłodziński M.: *Czy firmy świadczące usługi profesjonalne muszą tracić wartościowych klientów?* Marketing usług profesjonalnych. Kultura organizacyjna – osobowość profesjonalisty. Pod red. K. Rogozińskiego. Materiały z III-ciej międzynarodowej konferencji: 20-21 maj 2001; Poznań 2001.

CUSTOMER PROFILE IN CONSUMER MANAGEMENT RELATIONSHIP

Summary

Creating customer profiles consists of dividing a given group of customers into smaller categories on the basis of their common shopping habits and characteristics. Demographic and behavioral data are used to create an auxiliary picture of the client. The following text presents a profile of a client and a method of creating a customer profile in a consumer relationship management. The following text describes a profile of a client and a method of creating a customer. The described customer profile reflects demographic, personal data, customer preferences, likings and also the customer's value to the company. Personalization of the customer is beneficial to both parties. To the customer it means easier and faster access to the information, faster service and an individualized offer. To the company it means becoming a customer friendly organization which services its customers in a personalized manner. Also knowing its customer's preferences the company can create specific marketing and loyalty programs.

Sławomir Gawroński

**Katedra Dziennikarstwa i Komunikacji Społecznej
Wyższa Szkoła Informatyki i Zarządzania w Rzeszowie**

e-PR – PERSPEKTYWY PUBLIC RELATIONS W DOBIE SPOŁECZEŃSTWA INFORMACYJNEGO

Streszczenie

Rozwój technologii informatycznych, komputeryzacja codzienności i stale rosnące zainteresowanie Internetem w zasadniczy sposób wpływają na zmianę niemal wszystkich aspektów życia współczesnego człowieka. W szybkim tempie zmieniają się zasady komunikowania, zmienia się standard edukacji, przeobrażeniom ulega tradycyjny model pracy. Pod wpływem rozwoju technologii informatycznych i dochodzenia do modelu społeczeństwa informacyjnego, przeobrażeniom ulega również współczesna działalność marketingowa. O ile jednak w przypadku poszczególnych elementów marketingu-mix Internet pełni rolę wygodnego narzędzia wykorzystywanego w celu skuteczniejszej realizacji dotychczas realizowanych działań marketingowych, o tyle wobec public relations jest także czynnikiem stymulującym rozwój i ewolucję całej dziedziny, jaką jest PR.

1. WPROWADZENIE

Różnorodne determinanty rodzącego się społeczeństwa informacyjnego w zasadniczy sposób wpływają na zmianę niemal wszystkich aspektów życia współczesnego człowieka. Zmienia się niemal wszystko co nas otacza, a Internet, cyberprzestrzeń i wirtualna rzeczywistość stają się nieodłącznym elementem ludzkiego funkcjonowania. Pod wpływem rozwoju technologii informatycznych i dochodzenia do modelu społeczeństwa informacyjnego, przeobrażeniom ulega również współczesna działalność marketingowa. Jest kwestią oczywistą, że prowadzenie działalności marketingowej bez wykorzystania instrumentarium, jakiego dostarczają nowe technologie, jest jeśli nie niemożliwe, to przynajmniej mało efektywne. Coraz częściej pojawiają się więc terminy marketingowe, nawiązujące do nomenklatury „około internetowej”. Mówimy już zatem o „wirtualnym marketingu” [1], „e-marketingu” [2] czy też o „marketingu on-line” i „wirtualnej promocji”, co powoduje że swej wirtualnej „wersji” doczekało się także public relations. O ile jednak w przypadku poszczególnych elementów marketingu-mix Internet pełni rolę wygodnego narzędzia wykorzystywanego w celu stymulowania i skuteczniejszej realizacji dotychczasowych działań marketingowych, o tyle wobec public relations jest także czynnikiem stymulującym rozwój i ewolucję całej dziedziny, jaką jest PR.

Fraser P. Seitel wyróżnia pięć podstawowych czynników wpływających na rozwój public relations. Wśród nich wskazuje na rozwój Internetu, jako czynnika umożliwiającego

milionom ludzi na świecie dostęp do nieograniczonej czasem i przestrzenią komunikacji. „*Niesamowity rozwój Internetu i Sieci WWW szybko przeniósł public relations w XXI wiek. Niezliczona liczba użytkowników na całym świecie oznacza, że ludzie jednoczą się za pomocą komunikacji jak nigdy przedtem. Zmiany, jakie wnosi to zjawisko do społeczeństwa, są ogromne*”. [3]

Według badań sieci firm badawczych Taylor Nelson Sofres (TNS), w 2002 roku średnio 34% dorosłych użytkowników na świecie korzystało z Internetu (dane z 37 krajów). Oznacza to istotny wzrost w stosunku do lat ubiegłych (31% w 2001 roku i 27% w 2000) [4]. W Polsce w 2002 roku korzystanie z sieci zadeklarowało 18% dorosłych obywateli, nie mniej jednak, mimo rosnącego wskaźnika penetracji (wzrost o 3 punkty procentowe w stosunku do 2001 roku) nasz kraj w dalszym ciągu znajduje się w grupie państw o najniższym (poniżej 20%) wskaźniku dostępności do Internetu.

Rosnące znaczenie Internetu potwierdzają także badania amerykańskiego instytutu Harris Interactive dotyczące określenia najistotniejszych źródeł informacji wykorzystywanych przez Amerykanów [5]. Wynika z nich, że w dalszym ciągu media tradycyjne są najistotniejszym źródłem informacji, nie mniej znaczenie Internetu pod tym względem wciąż wzrasta. 76% ankietowanych jako najważniejsze źródło określiło telewizję, zaś po 8% respondentów wskazało radio i Internet. Zarówno telewizja jak i radio straciły na znaczeniu w stosunku do wcześniejszego badania przeprowadzonego przez tę samą firmę (78% – telewizja i 15% – radio), zyskał jedynie Internet (wzrost z 3% do 8%).

2. INTERNET A DZIENNIKARZE

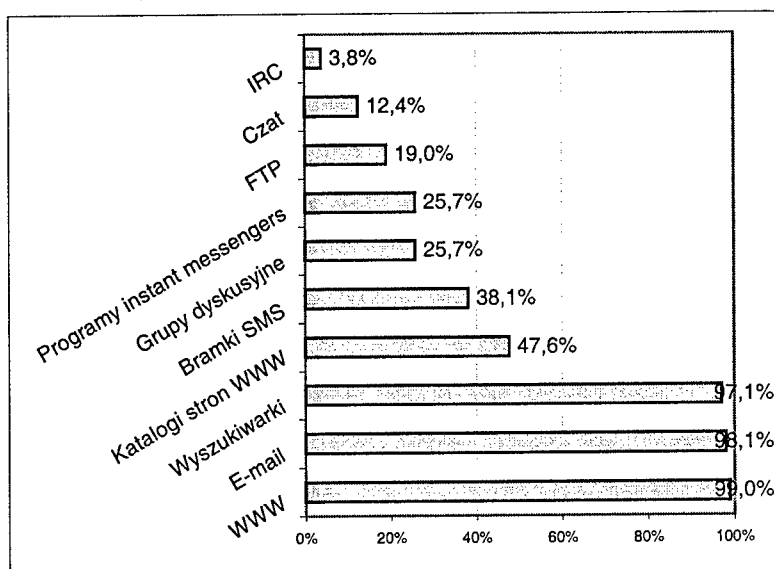
Internet łączy w sobie elementy medium skierowanego do odbiorcy masowego, jak i kanału komunikacji bezpośredniej. Innymi słowy, pozwala on na dystrybuowanie zarówno informacji skierowanej do masowego, nieznanego odbiorcy, jak i przekazu skierowanego do konkretnego użytkownika sieci. Dzięki interaktywnemu charakterowi Internetu, proces komunikacyjny przebiega w nim inaczej niż ma to miejsce w mediach tradycyjnych. Wykorzystując Sieć praktycznie każdy może być zarówno nadawcą, jak i odbiorcą przekazu komunikacyjnego. Specyfika Internetu powoduje także to, że odbiorca przekazów uzyskanych tym kanałem traci w pewnym sensie status odbiorcy masowego, gdyż tak naprawdę indywidualnie korzysta z kanału komunikacyjnego i sam decyduje o wyborze interesujących go informacji.

Powyższe tendencje, w odniesieniu do działalności marketingowej, podsumowuje T. Vassos: „Internet nie dotyczy masowego marketingu, ani masowego rynku. Dotyczy on ludzi - indywidualności z ich niepowtarzalnymi aspiracjami, potrzebami, pragnieniami i ich kulturowym tłem. Nie ma już masowego rynku, na przykład 60 milionów ludzi. Istnieje 60 milionów pojedynczych rynków, a każdy z nich liczy tylko jedną osobę” [6]. Jak zauważa Andrzej Świątecki, Internet ma podwójny charakter i pełni dualną funkcję. Jest bowiem zarówno narzędziem komunikowania marketingowego, jak i samodzielnym medium [7]. Znaczenie Internetu w komunikowaniu masowym systematycznie rośnie, co znacząco wpływa także na wykorzystanie go w działaniach public relations. Szczególnie istotne znaczenie odgrywa on w realizacji oddziaływań media relations. Znaczenie to jest wzmacniane również tym, że Internet stał się głównym narzędziem pracy polskich dziennikarzy. Internet jest bowiem źródłem informacji, najczęściej przez nich wykorzystywanym – wynika z badania dziennikarzy przeprowadzonego przez agencję public relations BCA S.A. i agencję badawczą ARC Rynek i Opinia [8]. Badanie dostarczyło zaskakujących wyników. Okazuje się bowiem, że pomimo tego, iż dziennikarze wykorzystują Internet zaledwie od 3-5 lat, aż 99% z nich deklaruje obecne korzystanie z Internetu, jako źródła informacji. Dodatkowo jedna trzecia ankietowanych dziennikarzy wskazała zasoby inter-

netowe jako najważniejsze źródło informacyjne dla przygotowywanych przez siebie publikacji i materiałów dziennikarskich.

O tym, że dziennikarze wysoko cenią Internet jako narzędzie wykorzystywane w swojej pracy świadczy również fakt, że niemal 90% uczestników badania wyraziło opinię, że ich praca bez Internetu byłaby trudna. Badanie po części obaliło mit Internetu, jako niewiarygodnego źródła informacji. Środowisko dziennikarzy, z racji wykonywanej profesji wyczulone na kwestię wiarygodności, w zdecydowanej większości (63,8%) oceniło to źródło informacji jako wiarygodne. Jedynie około 20% badanych wyraziło zastrzeżenia co do wiarygodności Sieci stwierdzając przy tym, że inne media są bardziej wiarygodnymi źródłami informacji.

Rozwój technologii informatycznej i upowszechnienie Internetu spowodowało gwałtowny wzrost liczby narzędzi elektronicznych wykorzystywanych przez dziennikarzy. Z badań przeprowadzonych na grupie polskich przedstawicieli tej profesji w 2001 roku wynika, że najpopularniejsze są instrumenty służące do komunikowania się (e-mail, bramki SMS, czat i programy typu instant messengers) oraz wyszukiwania informacji (wyszukiwarki, katalogi stron WWW, serwisy WWW i inne). Wyniki badań na temat form wykorzystywania Internetu przez polskich dziennikarzy przedstawia wykres 1.



Wykres 1. Wykorzystanie narzędzi internetowych w pracy dziennikarzy

Źródło: Opracowanie własne na podstawie „Internet w pracy dziennikarzy. Raport z badań ilościowych”, Warszawa 2001, <http://www.bca.com.pl>

Jednocześnie pojawia się zupełnie nowe spektrum wykorzystywania Internetu w relacjach pomiędzy specjalistami public relations a dziennikarzami, w ramach ich współpracy w zakresie działalności media relations. Coraz większa grupa dziennikarzy deklaruje wykorzystywanie Sieci jako sposobu przeprowadzania wywiadów, za pomocą internetowych narzędzi komunikacyjnych. Rośnie również odsetek dziennikarzy i specjalistów public relations współpracujących ze sobą, z wykorzystaniem najnowocześniejszych i nie spotykanych dotąd form realizacji działań media relations, jakimi są internetowe biura prasowe (on-line news bureau) oraz konferencje prasowe on-line. Udział w tych ostatnich deklaruje już 20% badanych dziennikarzy, a ponad 30% jest zainteresowana taką formą współpracy ze specjalistami w zakresie public relations.

3. INTERNET A PRZEDSIĘBIORCY

Ponieważ działalność media relations opiera się na wzajemnym komunikowaniu dziennikarzy z rzecznikami prasowymi i specjalistami w zakresie PR, zainteresowanie Internetem jako źródłem pozyskiwania informacji ze strony dziennikarzy ma swoje przełożenie na wykorzystywanie nowych technologii również w pracy PR-managerów. Internet na stałe zagościł w instrumentarium public relations wykorzystywanym przez przedsiębiorstwa, organizacje i instytucje publiczne. Stosowany jest powszechnie w praktyce PR nie tylko w obrębie sfery zadaniowej jaką jest media relations. Poczta elektroniczna (e-mail), sieci wewnętrzne (intranet), wewnętrzne wydawnictwa elektroniczne, a nawet wewnątrzfirmowe strony WWW coraz intensywniej wykorzystuje się również w ramach wewnętrznej komunikacji przedsiębiorstw. Internet w znacznym stopniu ułatwia także monitorowanie mediów on-line, co usprawnia i przyspiesza monitoring informacji i komentarzy na temat organizacji.

Public relations w wirtualnym wydaniu od pewnego czasu przestaje być domeną wyłącznie dużych przedsiębiorstw o zagranicznym rodowodzie, funkcjonujących na globalnym rynku. Badanie „Analiza sfer zadaniowych public relations przedsiębiorstw” [9] zrealizowane na obszarze badawczym województwa podkarpackiego wykazało, że Internet jest powszechnie wykorzystywanym środkiem wspomagania funkcji PR także wśród małych i średnich firm, funkcjonujących na lokalnych i regionalnych rynkach. Jak wynika z przeprowadzonych badań olbrzymia większość ankietowanych przedsiębiorstw wykorzystuje Internet w prowadzonej przez siebie działalności marketingowej, w tym w zakresie PR. Niemal 84% badanych firm świadomie i celowo wspomaga realizację swoich działań instrumentami internetowymi.

Korzystanie z Internetu jest w dużej mierze uzależnione od wielkości firmy. Badania potwierdziły, w pewnym sensie oczywistą tezę, mówiącą o wzajemnym powiązaniu wielkości organizacji z jej możliwościami wykorzystania nowoczesnych środków komunikowania. Zależność tę przedstawiono w tabeli 1.

Tabela 1

Korzystanie z Internetu a wielkość badanych przedsiębiorstw

		Liczba zatrudnionych w przedsiębiorstwie				Ogółem
		0-10 osób	11-50 osób	51-250 osób	Powyżej 250 osób	
Czy przedsiębiorstwo korzysta z Internetu?	Tak	67,9%	89,3%	100,0%	96,0%	83,7%
	Nie	32,1%	10,7%	0,0%	4,0%	16,3%

Źródło: Opracowanie własne na podstawie badań „Analiza sfer zadaniowych PR w przedsiębiorstwach”

Jak wynika z przedstawionych powyżej danych wzrost wielkości przedsiębiorstwa, mierzonej wielkością zatrudnienia w firmie, implikuje wzrost zainteresowania Internetem i stopnia jego wykorzystania. W grupie firm zatrudniających do dziesięciu osób odsetek przedsiębiorstw wykorzystujących Internet w swojej działalności wynosi 67,9%, a w grupie firm małych (zatrudniających od 11 do 50 pracowników) – już 89,3%. W firmach średnich i dużych wskaźnik wykorzystania Internetu w działalności przedsiębiorstw jest niemal stuprocentowy. A zatem średnie firmy w znacznie większym stopniu dostrzegają zalety Internetu, a w przypadku firm dużych korzystanie z sieci jest niemal obligatoryjne.

4. INTERNET A PUBLIC RELATIONS

Na fakt, że Internet staje się coraz istotniejszym narzędziem wykorzystywanym w ramach public relations wskazują dane dotyczące związku wykorzystania go w zależności od

deklarowanej istotności i znaczenia PR w badanych przedsiębiorstwach. Informacje na ten temat zawiera tabela 2.

Tabela 2

Korzystanie z Internetu a znaczenie działań PR

		Jakie znaczenie dla przedsiębiorstwa mają działania Public relations?				Ogółem
		Bardzo istotne	Dość istotne	Mało istotne	Nieistotne	
Czy przedsiębiorstwo korzysta z Internetu?	Tak	95,7%	89,6%	65,5%	57,1%	83,9%
	Nie	4,3%	10,4%	34,5%	42,9%	16,1%

Źródło: Opracowanie własne na podstawie badań „Analiza sfer zadaniowych PR w przedsiębiorstwach”

Z danych przedstawionych w tabeli 2 w wyraźny sposób wynika, że wzrost znaczenia działań public relations w przedsiębiorstwie pociąga za sobą zainteresowanie możliwościami, jakie daje Internet w sferze działalności marketingowej przedsiębiorstwa. Wśród badanych przedsiębiorstw, w których prowadzone działania PR uznawane są za nieistotne, odsetek firm wykorzystujących Internet w swojej działalności wynosi 57,1%. Wraz ze wzrostem znaczenia PR, rosną także wskaźniki mówiące o powszechności wykorzystania Internetu. Odsetek firm korzystających z Sieci wynosi bowiem 65,5% dla przedsiębiorstw oceniających swoją działalność PR jako mało istotną, 89,6% – w grupie firm dla których oddziaływanie public relations są dość istotne i aż 95,7% – dla organizacji najwyższej oceniających znaczenie PR. Wykorzystanie Internetu wśród praktyków PR wzrasta zatem wraz ze wzrostem świadomości menedżerów odnośnie potrzeby prowadzenia tego typu działalności i znaczeniem działań PR w całym przedsiębiorstwie. Po raz kolejny potwierdza to tezę, że menedżerowie realizujące zadania public relations w sposób celowy, świadomy i z wiarą w ich skuteczność, wykorzystują całość instrumentarium PR, jakie mają do dyspozycji, a tym samym wkraczają ze swoją aktywnością w coraz to nowe sfery zadaniowe public relations.

Tabela 3

Sposoby wykorzystania Internetu w działaniach PR badanych przedsiębiorstw

Sposób wykorzystania Internetu	Procent przypadków
Strona WWW, witryna internetowa przedsiębiorstwa	80,6%
Mailing bezpośredni	49,7%
Grupy dyskusyjne i listy adresowe	14,5%
Biuletyny elektroniczne	14,5%
Subskrypcja informacji (newsletter)	9,7%
Internetowe biuro prasowe (on-line news bureau)	3,6%
Ogółem	172,6%

Źródło: Opracowanie własne na podstawie badań „Analiza sfer zadaniowych PR w przedsiębiorstwach”

Z danych przedstawionych w tabeli nr 3 wynika, że najczęściej stosowaną formą wykorzystania Internetu w badanych przedsiębiorstwach jest posiadanie własnej witryny internetowej. Własną stronę internetową posiada aż 80,6% badanych przedsiębiorstw korzystających z Internetu. Strona wykorzystywana jest przeważnie jako nośnik informacji o produktach, usługach, a także jako zbiór danych teleadresowych. Inną formą chętnie stosowaną w badanych firmach jest mailing bezpośredni (49,7% wskazań), a także grupy dyskusyjne i listy adresowe (14,5% wskazań) oraz biuletyny elektroniczne (14,5% wskazań). Badane przedsiębiorstwa rzadziej wykorzystują Internet do subskrypcji informacji

(9,7% wskazań), a jedynie 3,6% badanych firm wykorzystuje Internet do prowadzenie biura prasowego on-line.

Przedstawione dane wskazują na bardzo tradycyjne wykorzystywanie możliwości, jakie daje Internet w zakresie prowadzenia działań PR. Wynika to ze specyfiki firm funkcjonujących na badanym obszarze. Nowoczesne (ale i dość drogie) instrumenty internetowe (on-line news bureau, serwisy fotograficzne on-line, etc.) pozostają wprawdzie w dalszym ciągu poza możliwościami większości badanych przedsiębiorstw, ale zainteresowanie nimi stale wzrasta, co może być zwiastunem swego rodzaju rewolucji internetowej w zakresie public relations.

5. WNIOSKI KOŃCOWE

Przedstawione wyniki badań środowiska dziennikarzy i specjalistów w zakresie public relations dotyczące stopnia i zakresu wykorzystywania przez nich Internetu pozwalają na potwierdzenie tez, stawianych od kilku lat przez teoretyków zagadnienia, mówiących o rosnącej sile oddziaływania technologii informatycznych i nowych mediów opartych na sieci komputerowej w komunikowaniu public relations. Zrewolucjonizowanie niemal wszystkich dziedzin ludzkiego życia przez efekty tworzącego się społeczeństwa informacyjnego w głównej mierze dotyczy sfery komunikowania. Public relations, jako istotna gałąź komunikacji społecznej, również przenikane jest konsekwencjami rewolucji komunikacyjnej. Przewidywania związane z rozwojem tej dziedziny potwierdzone obserwacjami rosnącego wpływu Internetu na zachowania komunikacyjne public relations pozwalają przypuszczać, że „wirtualne PR”, jako forma nowocześniejsza i skuteczniejsza od tradycyjnego modelu, będzie nadal zwiększało swoje znaczenie.

BIBLIOGRAFIA

- [1] A. Sznajder, *Marketing wirtualny*, Warszawa 2000
- [2] M. Kacprzyk, „Świat według e-”, „Marketing Serwis”, Nr 12/2000, s. 64-67
- [3] F. P. Seitel, *Public Relations w praktyce*, Warszawa 2003, s. 26
- [4] Taylor Nelson Sofres, *Światowy raport o Internecie*, lipiec 2002, <http://www.obop.pl>
- [5] Wyniki badania ze strony internetowej
<http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=371>
- [6] T. Vassos, *Strategic Internet Marketing*, London 1996, s. 116
- [7] A. Świątecki, *Zintegrowane Komunikowanie Marketingowe*, Warszawa 2001, s. 112
- [8] „Internet w pracy dziennikarzy. Raport z badań ilościowych”, Warszawa 2001,
<http://www.bca.com.pl>
- [9] D. Tworzydło, G. Humenny, R. Góra, S. Gawroński, T. Soliński, *Analiza sfer zadaniowych public relations w przedsiębiorstwach*, Instytut Gospodarki, Rzeszów 2003

e-PR – PUBLIC RELATIONS' PERSPECTIVES IN THE INFORMATION SOCIETY

Summary

The development of information technologies, computers and Internet are changing nearly all elements in our modern life. The communications rules, educational standard and our work style are changing fast. These modifications are leading us to information society and it has changed marketing rules too. Internet is not only a helpful tool, used by all elements of marketing mix but it is an important factor, that stimulates and causes a really big evolution of public relations activities.

Marek Hetmański

Instytut Filozofii UMCS, Lublin

JAKIEJ INFORMACJI CZŁOWIEK POTRZEBUJE NAPRAWDĘ

Streszczenie

W świecie masowej komunikacji, w multimedialnym środowisku, gdzie dominuje zmysłowy obraz a niedoceniane jest słowo, człowiek odbiera narastającą lawinę sygnałów. Trudno jest je selekcjonować czy oddzielić od szumów. Same sygnały nie są w pełni informacją, jakiej człowiek potrzebuje w poznaniu czy do wiedzy, którą chce mieć o świecie. Informacją jest dopiero ten sygnał, szerzej, komunikat, który steruje działaniem człowieka, zmniejsza jego niepewność czy niewiedzę, który przyczynia się do powstania nowej wiedzy. Informacją jest zatem ten sygnał, który w obrębie procesu komunikacyjnego ma znaczenie, jest przez jego uczestników rozumiany, interpretowany, który również może i musi być przedmiotem osądu, oceny, a także odrzucenia. Nowoczesne technologie komunikacyjne nastawione są na produkcję i przekaz sygnałów, które w pewnych tylko sytuacjach poznawczych, w niewielkim zresztą stopniu, mają wartość informacyjną. Nadzieje na budowanie w oparciu o samą tylko infrastrukturę informacyjną przyszłego społeczeństwa wiedzy są wciąż złudne, nie wiadomo czy się spełnią. Stawiają bowiem człowieka w obliczu samych sygnałów, a nie znaczących komunikatów. Z najbogatszego nawet zestawu multimedialnych znaków nie wyłoni się samoczynnie wieloznaczny komunikat czy symbol. A tylko takie komunikaty (im bardziej niejednoznaczne i mniej prawdopodobne, tym więcej zawierające informacji) wzbogacają ludzką wiedzę. Póki co żyjemy w świecie coraz bardziej jednoznacznym i globalnym, paradoksalnie, coraz bardziej pozbawionym... informacji.

1. WSTĘP

Człowiek nie potrzebuje w swoich czynnościach praktyczno-poznawczych informacji jako takiej, informacji samej w sobie. Zasadniczo potrzebna jest mu wiedza, ale też nie wiedza w ogóle nie wiedza wszelka; taka jest tylko epistemologicznym ideałem nieosiągalnym na co dzień. Człowiek potrzebuje wiedzy określonej ze względu na pewne sytuacje poznawcze, potrzeby, wyznaczone sobie cele. Poszukując wiedzy, zdobywając jej pewne rodzaje, przechodząc od wiedzy mniej pewnej do wiedzy o pewności większej, ale też i odwrotnie, człowiek za każdym razem percypuje, zapamiętuje, przetwarza i przekazuje informacje. Nie ma zatem wiedzy bez informacji, lecz tym, o co naprawdę w poznaniu idzie jest znacząca wiedza, a nie sygnały charakteryzujące się określoną miarą informacji, które tylko w pewnych sytuacjach przyczyniają się do powstania wiedzy.

Jaka zatem informacja ma wartość dla powstania wiedzy? Jakie sytuacje praktyczno-poznawcze warunkują i sprzyjają wiedzy ludzkiej? Na te *stricte* epistemologiczne pytania informatyka nie udziela odpowiedzi. Informatyka jest bowiem nauką o sygnałach i ich materialnych nośnikach, o regułach przekształcania znaków, o programach, narzędziach i systemach informatycznych, w tym sensie nie zajmuje się ona wiedzą a tylko informacją. Epistemologia z drugiej strony, przynajmniej ta tradycyjna, nie zauważa czy też nie docenia na ogół faktu technicznego wsparcia procesów poznawczych, tego że mają związek z określoną miarą informacyjną sygnałów i komunikatów. Zajmuje się na ogół naturalnymi czynnościami poznania, zazwyczaj też wyidealizowanymi, ustala normy poznawcze. Co zatem epistemologia może powiedzieć informatyce na temat natury informacji? Co z kolei informatyka ma ważnego do powiedzenia o wiedzy, z czym filozofia mogłaby się liczyć? Kwestia ta nie jest wyłącznie akademickiej natury, gdyż dotyczy spraw praktycznych – tego jak projektować systemy informatyczne, szczególnie ich interfejsy z człowiekiem, aby mógł on je efektywnie obsługiwać, korzystać z nich nie tylko jako ze źródła samej informacji, ale również mieć dzięki nim szansę zdobycia znaczącej wiedzy. Idzie zatem o bardziej ludzkie, podmiotowe oblicze informatyki, jak również o pogłębioną analizę zależności między informacją a wiedzą.

2. CZY SYGNAŁY SĄ INFORMACJĄ

Nie od rzeczy jest przypomnieć (nawet w obecności informatyków, a może przede wszystkim dla nich), czym w ogóle jest informacja, jak ona funkcjonuje w świecie fizycznym czy w technice, jak pojawia się w świecie ludzkich działań praktycznych i poznawczych i jak przekształca się czasami, bo nie zawsze i nie bezproblemowo, w wiedzę.

Jako istoty żywe i wyposażone w zmysły żyjemy w świecie fizycznych bodźców, które ze względu na sensoryczną wrażliwość naszego ciała stają się sygnałami dla naszego działania. Przebiega tak działanie zarówno na elementarnym poziomie fizjologicznych reakcji wewnętrznych (np. pracy układu nerwowego), jak i zachowanie się organizmu w środowisku tak przyrodniczym, jak i społecznym czy kulturowym. Różnorodne bodźce są w trakcie percepcji przekształcane w jednolite impulsy nerwowe, które stają się następnie podstawą pojawiania się w umyśle różnorodnych wrażeń i spostrzeżeń. Ten wieloetapowy proces ma charakter kodowania i dekodowania sygnałów, jest więc procesem *stricte* informacyjnym. Nie na nim jednak kończy się proces poznania, nie tutaj też tworzy się od razu wiedza, to zaledwie początek złożonego mechanizmu wiedzytwórczego. Koncepcje człowieka jako procesora informacji, które powstają w nurcie tzw. sztucznej inteligencji, poprzestające w analizach poznania na tym etapie, nie mówią w ogóle o wiedzy i dają fałszywy obraz człowieka.

Nie jesteśmy przepastnym pojemnikiem na sygnały, nie jesteśmy magazynem dla wszelkiej i dowolnej informacji, nasz umysł nie jest filmową błoną rejestrującą wszystkie zewnętrzne bodźce. Istnieją zarówno w naszym ciele, jak i w naszych narzędziach poznawczych, zwyczajach czy instytucjach społecznych liczne i różnorodne progi i bariery wrażliwości na sygnały, istnieją też granice absorpcji sygnałów. Jesteśmy istotami przyswajającymi i selekcionującymi informacje (*we are what we behold*, jak mówi M. McLuhan), poza które nieustannie wykraczamy (*beyond information given*, jak z kolei mówi J. Brunner). Daleko tym samym sięgają obszary naszego poznania, które polega nie tylko na bezpośrednim związku z doznaniem zmysłowymi. Dlatego też wiedza to, coś więcej niż dostarczone i przetworzone informacje. Naszym życiem umysłowym rządzą

mechanizmy rozróżniania, selekcji, wyboru i oceny tego, co widzimy, słyszymy, dotykamy tego, co bezpośrednio poznajemy. Nie każdy bodziec jest jednakowo ważnym sygnałem i nie każdy sygnał jest zawsze i tak samo przez organizm wykorzystywany. Ten fizjologiczny mechanizm decyduje o różnej wartości recypowanych sygnałów, przydając im różną wartość informacyjną dla ciała. Podobne mechanizmy działają także na poziomie świadomej reakcji człowieka na widziane rzeczy czy przeżywane stany umysłowe. Zmysłowe wrażenia i umysłowe obrazy (nazywane przedstawieniami), powstałe poprzez oddziaływanie i odbieranie impulsów i sygnałów - stanowią zawartość naszego umysłu, naszą indywidualną, bogatą i różnorodną treść doświadczenia - same też są sygnałami. Są to sygnały z wyższego poziomu doświadczenia, które stają się przyczyną i podstawą szerszych oraz bardziej złożonych jednostek wiedzy, takich jak przypomnienia, wyobrażenia, pojęcia, wyrażenia językowe, fantazje. One też mają swoją wartość informacyjną dla działania człowieka, gdy kieruje się on nimi jako motywami czy celami swojego jednostkowego działania. Ale też i one są z kolei sygnałami do dalszego, bardziej już złożonego działania.

Znaczenie wyobrażeń, pojęć czy fantazji, nade wszystko zaś wyrażen językowych jak słowa, pisane znaki, zdania i wypowiedzi mowy potocznej czy też języka naukowego (które mają jednoznacznie informacyjny charakter), zasadniczo wykracza poza jednostkowe doświadczenie. Jest to doświadczenie ponadindywidualne – grupowe, zbiorowe, społeczne, także ogólnoludzkie. Poszczególne składowe tego ogólnego doświadczenia występują także w roli sygnałów, które każdy z nas odbiera zwrótnie w swoim osobistym doświadczeniu – postrzegamy i reagujemy na uniwersalne znaki i symbole, rozwiązujemy pewne sytuacje problemowe (codziennego życia czy nauki), odpowiadamy na konkretne pytania ale i uniwersalne wezwania, jakie stawia przed nami cywilizacja i kultura. Za każdym razem bezpośrednio doznajemy sygnałów, lecz ich wartość informacyjna jest już inna niż w przypadku pojedynczych wrażeń i spostrzeżeń zmysłowych składających się na obrazy czy pamięć.

Zawartość informacyjna wieloznacznych znaków, symboli i wartości stanowiących treść komunikatów kulturowych nie ma bezpośredniego związku z materialnymi nośnikami sygnałów, kodowaniem znaków czy maszynowymi procesami przetwarzania informacji. Informacyjna treść tych jednostek wiedzy ogólnej nie łączy się już bezpośrednio (choć nadal ma związek) z ich sterującą rolą w życiu jednostki, lecz zasadniczo ze znaczeniem i sensem, jakie mają one dla wiedzy. Słowem, symbole-sygnały, z jakimi mamy do czynienia w nauce czy kulturze odznaczają się inną zawartością informacyjną niż proste sygnały, z jakimi ma do czynienia na co dzień informatyka i jej coraz liczniejsze zastosowania. A ponieważ cywilizacja współczesna ulega coraz bardziej z informatyzowaniu („zdygitalizowaniu”, jak brzmi nieznośny slogan coraz częściej bezmyślnie powtarzany), refleksja nad tymi zjawiskami winna być prowadzona w samej informatyce, lecz przy wsparciu ze strony epistemologii.

3. ZMIANY TEGO, CO WIDZIALNE

Tym, co informatyka winna uwzględnić w swojej teorii (poszukiwaniu nowych reguł kodowania czy programowania), jak również w praktyce (projektowaniu i budowaniu systemów informatycznych) jest fakt zasadniczy – powstanie nowych reguł percepcji i organizacji wiedzy o świecie. Na skutek zmian spowodowanych przez technologie informatyczne (o czym piszą m.in. M. Minsky, H. Simon, N. Negroponte, M. Castells) powstała cywilizacja i kultura charakteryzujące się nową wizją człowieka i świata

(o czym z kolei piszą m.in. S. Turkle, S. Papert, D. Bolter, D. deKerckhove, J. Baudrillard, P. Virilio, M. Heim). Techniki cyfrowego zapisu dowolnych sygnałów w postaci określonych danych ułatwiły ich wizualizację. W zasadzie człowiek zawsze obrazował to, co bezpośrednio postrzegał, dokonując tego w zależności od techniki, jaką w danym czasie dysponował. Technika cyfrowego kodowania, która uzupełniła zapis analogowy, wcale go nie wypierając, przyczyniła się do nowej wizualizacji świata. W miejsce dotychczasowych obrazów rejestrujących bezpośrednio doznania i ich przetworzone (np. w sztuce czy w nauce) obrazy zaczęły pojawiać się obrazy o zupełnie nowej jakości. Obrazy bezpośredniej lub przetworzonej sfery doznań zmysłowych zasadniczo reprezentowały i odzwierciedlały (zgodnie z klasyczną teorią prawdy jako odsłaniania rzeczywistości) świat realny i fizyczny. Przedstawiały to, co było zakryte przed wzrokiem, a więc zbyt odległe czy za małe, lecz z natury rzeczy dające się w ogóle poznać. Ustanowiły specyficzny epistemologiczny paradygmat widzialności – percepcji i myślenia o tym, co jest realne.

Z paradygmatem takiej widzialności miała do czynienia od początku swojej historii informatyka. Podjęła się zadania kodowania i obrazowania dyskretnych sygnałów oraz ich, w pierw w zdaniowej, a potem graficznej, wizualizacji. Dotychczasowe analogowe obrazy zaczęły zastępować obrazy o dyskretnym sposobie kodowania. Informatyka przyjęła zatem ideę widzialności jako takiej w jej pierwotnej postaci – bezpośredniego, bezproblemowego i adekwatnego (prawdziwego) widzenia świata. Ale technologie informatyczne pozwoliły także na coś więcej – na taką wizualizację danych, która stworzyła iluzję kreowania światów nieistniejących, zupełnie wyimaginowanych, wirtualnych. Wirtualna rzeczywistość jest produktem nowego paradygmatu widzialności – prezentacji bardzo bogatych w doznania zmysłowe (wykorzystujących efekty synergii i złudzeń zmysłowych) obrazów szczególnych aspektów rzeczywistości, w szczególności nierealnych powiązań czy relacji pomiędzy jej fragmentami. Zerwanie z tradycyjną optyką i sprowadzenie poznania do wirtualizacji rzeczywistości a nie jej odzwierciedlania było możliwe dzięki specyficznej „maszynizacji wzroku”. Przełamana została dominująca w poprzednim paradygmacie centralność oka w widzeniu, odrzucone zostało wyobrażenie percepcji i umysłu jako odzwierciedlania zewnętrznej rzeczywistości. Nastąpiła „epoka wizjoniki”, według określenia P. Virilio, lub też „epoka wideosfery”, nazwana tak z kolei przez R. Debraya. Człowiek jako użytkownik wyrafinowanych narzędzi informatycznych i technik audiowizualnych zaczął osiągać szczególne widzenie bez patrzenia. W zakres nowej widzialności zaczynają wchodzić obszary świata, których nigdy nie było i nie będzie, choć są one prezentowane w niespotykanej dotychczas intensywności obrazów wykreowanych przez nowe media.

4. KIEDY INFORMACJA STAJE SIĘ WIEDZĄ

Powyższych zmian w poznaniu świata i ich cywilizacyjnych skutków nie należy przypisywać wyłącznie informatyce. Wprawdzie zapoczątkowała ona proces cyfrowej rewolucji w przetwarzaniu informacji, to jednak nowy paradygmat współtworzyły (wcześniejsze) zmiany w systemach telewizji, kina i masowej kultury. To one wprowadziły nas w świat multimedialnych obrazów, które zaczęły dominować nad słowem mówionym i tekstem pisanym, zmieniając nasze poznanie, w którym intensywność doznań zmysłowych, głównie wzrokowych, na plan dalszy odsunęła rozumienie i poszukiwanie sensu tego, co poznawane. Dlatego też priorytet zaczęły zyskiwać sygnały i proste przekazy o małej zawartości informacyjnej, zaś znaki i symbole właściwe złożonym i bogatym komunikatom, przesunięte na marginesy kultury masowej, przestały koncentrować naszą

uwagę. Nie jest to zamierzone działanie samej informatyki, lecz jedynie efekt uboczny rewolucji cyfrowej przez nią zainicjowanej. Niemniej informatyka winna uświadomić sobie rozmiar tej zmiany i jej skutki. Na nowo trzeba zanalizować i ocenić rolę informacji w różnorodnych systemach komunikacji. Ale takiej informatyki, która, skrótowo mówiąc, będzie uwzględniała potrzeby i możliwości użytkowników narzędzi i systemów informacyjnych, a nie wyłącznie ich projektantów, administratorów czy dystrybutorów. Informatyki wyzwolonej od mechanizmów marketingu, otwartej na sprawy publiczne, prowadzącej dyskusję nad kluczowymi dla niej pojęciami informacji i wiedzy. Wyjść zaś trzeba od kilku podstawowych ustaleń i twierdzeń, które są następujące:

Po pierwsze, sygnały, które odbiera człowiek mają wtedy dopiero wartość informacyjną, gdy sterują jego działaniem; nie wtedy jednak, gdy zaledwie je on przyswaja, wchłaniając informacje według mechanizmu sprzężenia zwrotnego dodatniego, lecz wówczas dopiero, gdy sprzężone są one z nim ujemnie, a więc gdy modyfikują (redukują nadmiar) działanie w jego trakcie. Ten bowiem sygnał ma znaczenie, który uczestniczy w swoistym „metabolizmie informacyjnym”.

Po drugie, sygnały i przekazy, które są nowe, niespotykane, zaskakujące – a więc mniej prawdopodobne w stosunku do wysokiego, czy co najmniej równego prawdopodobieństwa sygnałów i komunikatów już spotykanych i znanych – mają większą wartość informacyjną. Im komunikaty są mniej prawdopodobne, tym więcej niosą ze sobą informacji. Dla użytkownika systemów informacyjnych ma to duże znaczenie, gdyż tak naprawdę oczekuje on tego, co nowe i zaskakujące a nie powtarzalne i oczekiwane. Nie idzie wszak o zaskakiwanie (jak w przypadku multimedialnej rozrywki), lecz rozbudzanie w odbiorcy postaw poszukiwania tych ważnych wiadomości, które wykraczają poza horyzont jego dotychczasowej wiedzy. Jest to szczególnie ważne w dziedzinie edukacyjnych zastosowań informatyki.

Informacją jest zatem, po trzecie, ten sygnał i przekaz, które stawiają podmiot przed wyborem ich spośród innych sygnałów. Takim jednak wyborem, który dla podmiotu łączy się z koniecznością podjęcia decyzji o poznawczym charakterze. Wybór taki musi opierać się na umiejętności nie tylko absorpcji bodźców, lecz głównie ich selekcji, wartościowaniu, porównywaniu i w końcu redukowaniu.

Po czwarte zatem, i co najważniejsze, przekazy i sygnały muszą być odbierane w kontekście istniejącej już wiedzy. Kontekstowość informacji polega na nadawaniu przez nadawcę przekazom określonych znaczeń, zaś przez odbiorcę ich odnajdowaniu i ustanawianiu. Dla przekazów czysto zmysłowych będzie to odnoszenie ich do kontekstu pojęciowego, zaś dla przekazów pojęciowych do kontekstu obrazowego. Gra spostrzeżeń z pamięcią i wyobraźnią jest najlepszym warunkiem tworzenia się wiedzy w oparciu o informacje.

WHAT INFORMATION WE NEED IN FACT**Abstract**

Thanks to the mass communication we live now in the world full of signals, especially intensive visual pictures that are our immediate cognitive environment. But signals do not constitute our knowledge, they are even not the information. Only these signals and messages that control and regulate our practical and cognitive behavior (appropriately to our intentions) have informational significance. We need in substance the meaningful information and not the information as such. The human beings besides need the information that reduces the uncertainty of their knowledge and changes the probability of received messages. We are the subjects that absorb, select, and evaluate or reject the given information. Thanks to these mechanisms we have the meaningful knowledge. This epistemological thesis must be considered in the area of computer sciences and computer engineering that deal so far only with signals and information

Kazimierz Krzysztofek

Instytut Socjologii, Szkoła Wyższa Psychologii Społecznej w Warszawie

GLOBALNE SPOŁECZEŃSTWO INFORMACYJNE: TRZY TRENDY

Uczni od dawna głowia się nad tym, jaka będzie organizacja przyszłego społeczeństwa, jaki system regulacji. Generalnie można tu wyróżnić trzy podejścia: które opatrują kryptonimami Matrix, Arkadia i Deja vu.

1. SPOŁECZEŃSTWO TYPU MATRIKS

Matrix to najkrócej rosnąca technicyzacja ludzkiej egzystencji i unaukowienie przetrwania (*scientification of survival*), czyli uzależnienie człowieka od coraz to nowszych generacji techniki i odkryć naukowych oraz obawy przed zredukowaniem komunikowania do interakcji człowiek-maszyna.

Wielu jest przekonanych, że nowa regulacja może przybrać kształt technototalitarny i ludzie zostaną tak zdeterminowani przez cywilizację informatyczną, że nie będą w stanie się obronić. Temat ten staje się obsesją kultury popularnej (filmy typu Matrix, Equilibrium, AI, Robocop i wiele innych mogą być ostrzeżeniem, ale też samospełniającym się proroctwem). Kto w takim systemie się nie znajdzie, nie będzie zdolny do funkcjonowania w społeczeństwie, bo żadna emigracja wewnętrzna nie pomoże. Będziemy żyć pod nadzorem „elektronicznego pastucha”. Bycie w systemie to nie tylko korzystanie z jego dobrodziejstw, lecz również straty w sferze wolności, prywatności czy wręcz intymności.

W niektórych projekcjach ów „system” czy wręcz „metasystem” (system integrujący wszystkie inne) jest postrzegany jako inkorporacja biosocjiosfery przez biotechnosferę - widać tu powrót determinizmu (Chyła 2003). Najgłośniejszy wywołał ten problem Neil Postman w książce „Technopol. Triumf techniki nad kulturą” (1995), rychło znajdując zwolenników. W ich przekonaniu wszystkie technologiczne przedłużenia człowieka, interfejsy człowiek-maszyna, są wymyślane po to, aby jak najbardziej zintegrować go z systemem kontroli przez „wrastanie maszyny w nas samych”. Jak pisał C. Talbott w „*The Future does not compute*” (1995), choć ciągle nam się wydaje, że panujemy nad maszyną, to nie możemy uciec od prawdy, że ona zagnieżdża się w nas przez interfejs. Jak wielki jest to wpływ, pokazuje dyfuzja telefonii mobilnej (pokolenie SMS)

Wielu jest przekonanych, że ta nowa regulacja może przybrać kształt totalitarny i ludzie będą tak zdeterminowani przez cywilizację informatyczną, że nie będą w stanie się obronić. Kto w takim systemie się nie znajdzie nie będzie zdolny do funkcjonowania w społeczeństwie, żadna emigracja wewnętrzna nie pomoże.

„Być w systemie” to znaczy nie tylko korzystać z jego dobrodziejstw, ale też wiele utracić w sferze wolności, prywatności, czy wręcz intymności. W niektórych projekcjach widzi się ów „system”, czy wręcz „metasystem” (system integrujący wszystkie inne) jako inkorporację biosocjosfery przez biotechnosferę; widać tu więc powrót determinizmu. Im więcej owych przedłużeń tym więcej ewidencjonowania wszelkich przejawów życia. Nie da się już nawet policzyć, a przynajmniej obywatel tego nie wie, w ilu bazach danych się znajduje. Nie wiadomo, czy ziści się na koniec przyszłego wieku straszna wizja *machina sapiens*, czy też człowiekowi uda się wyzwolić jakieś antyciała na swą obronę. Zanim to nastąpi – jeśli w ogóle – problemem XXI wieku będzie to, z czym już dzisiaj mamy do czynienia, mianowicie z ukrytym sterowaniem ludźmi, tyle że w znacznie większym nasileniu.

Człowiek posługujący się coraz potężniejszymi narzędziami kieruje dziś swe zainteresowanie bardziej ku sobie i swemu wnętrzu niż ku światu. Była już kolonizacja świata, teraz mamy k(o)lonizację człowieka. „Wkroczyliśmy w epokę, stwierdza Zbigniew Brzeziński (1998: 10), w której nauki przyrodnicze przeobrażają się z narzędzia podboju środowiska zewnętrznego w narzędzie podboju – by tak rzec – wewnętrznego środowiska człowieka. Inaczej mówiąc, nauka przechodzi od podboju natury do podboju samej istoty ludzkiej. Dotychczasową historię człowieka można było opisać jako ciągły postęp w rozwijaniu naszej zdolności do rozumienia i kontrolowania zjawisk wobec nas zewnętrznych. Uprawa roli, przemysł, ekonomia, wyprawy kosmiczne – to kolejne kroki w osiągnięciu w zdolności do zapanowania przez ludzkość nad środowiskiem zewnętrznym. Dzisiaj natomiast najbardziej dramatyczne odkrycia naukowe odnoszą się w coraz większym stopniu do „kontroli wewnętrznej” – do tego, czym istota ludzka jest i czym może się stać”.

Czy takie społeczeństwo można jednak opatrywać szyldem złowrogiego Matrixa. Może jest on w każdym z nas?

2. SPOŁECZEŃSTWO SIECIOWE CZYLI ARKADIA

Istnieje silny nurt optymistyczny w ocenie dotychczasowych tendencji oraz w prognozowaniu rozwoju społeczeństwa informacyjnego. Wśród optymistów znaleźć można wybitnych badaczy, jak Georges Gilder, Ray Kurzweil, Arun Netravali, Nicholas Negroponte, Marvin Minsky, Hans Moravec, Eric. K. Drexler, nie żyjący już Michael L. Dertouzos którzy jednocześnie byli lub są mocno zaangażowani w rozwój technologii informacyjnych, a niektórzy są szefami wielkich laboratoriów, wylęgarni innowacji technologicznych i oczywiście zależy im na sukcesie swych przedsięwzięć. To wyjaśnia powody owego optymizmu. Ci którzy tak epatują nas wizją *machinae sapientis* – świadomie czy nieświadomie wykorzystują to, że o mózgu nadal niewiele wiemy i nie jesteśmy w stanie dziś zweryfikować tych wizji. Daleko w optymistycznych wizjach idzie ogłoszony w lipcu 2000 r. Raport Amerykańskiej Narodowej Fundacji Naukowej (<http://www.nsf.gov/od/1pa/news/02/pr0257.htm>), w którym mowa jest o misji biotechnologii i informatyki w dziele udoskonalenia, człowieka, aby zbudował lepszą cywilizację.

Tu już ujawnia się przewaga optymizmu, wskazuje się bowiem na nadzwyczajne szanse rozwoju (także kulturowego), zarówno dla społeczeństw już zaawansowanych cywilizacyjnie, jak i tych odstających, „przedinformatycznych”. Można się w tym dopatrywać nostalgii za lepszym światem, wiary w to, że cywilizacja ciągle się doskonali, przechodząc na coraz wyższe etapy swego rozwoju, który opatrywany pozytywnym znakiem - był synonimem postępu. Jest to wiara w swym rodowodzie oświeceniowa, która od dawna ulega erozji, ale której fale powracają. Pojawiają się poglądy, że po raz pierwszy

od „Galaktyki Gutenberga”, która narzuciła *homo sapiens* linearny (jak drukowany wiersz w książce czy w gazecie) paradygmat percepcji, kultura multimediów harmonizuje z naturalną multimedialnością człowieka, który pisze, mówi, tworzy obrazy i jeszcze do tego czuje. Może dzięki tej harmonizacji epoka cyfrowa doprowadzi do bezprecedensowego rozkwitu podmiotowej edukacji, a nie zaprogramowania człowieka.

Wizje optymistyczne bazują na wierze w uzdrawiającą moc sieci społecznych, dla których sieci komputerowe są znakomitą ekstensją. Francuski filozof sieci, Pierre Levy (1996) wierzy, że ITs będą mieć humanizujący wpływ; rozprasza on obawy, że maszyny będą nami rządzić. Jest wręcz przekonany, że swobodna wymiana idei w Internecie będzie nas stopniowo uwalniać od społecznych i politycznych hierarchii, które stoją na drodze postępu ludzkości.

Levy, i jego kanadyjski kolega, Derrick de Kerckhove (1999), na każdym kroku manifestują swą wiarę, że spotkanie umysłów w sieci prowadzi do wykształcania się kolektywnej inteligencji. Dzięki Sieci grupa ludzi może działać i wykonywać jakieś zadanie, jak gdyby była inteligentnym, samomonitorującym się organizmem pracującym z jednym umysłem, a nie zbiorem niezależnych „agentów intelektualnych”. Już dziś można mówić o czymś w rodzaju rozproszonej inteligencji – *dispersed processing* – połączone komputery, które niesamowicie multiplikują w sieci moc obliczeniową, np., przy analizie sygnałów pochodzących z obcych planet. W społeczeństwie sieciowym tożsamość człowieka nie jest związana już tylko z rodziną, wspólnotą lokalną czy państwem, ale także z globalną siecią. Pojedyncza osobowość staje się częścią oplatającej świat całości.

ITs jawią się jako technologie wolności. Jest to wizja społeczeństwa komunikującego się, którego socjologiczne charakterystyki opisał Juergen Habermas. Rodzi się nowa wspólnota komunikacyjna; intersubiektywność, kolektywna inteligencja prowadzą do wzajemnego porozumienia jednostek bez przymusu. Taka kultura rodzi nowe tożsamości. Charakteryzuje je nieskończona zdolność łączenia wszystkiego ze wszystkim i niesamowita zdolność wchłaniania kultury, ekonomii, polityki – wszystkiego. Społeczeństwa zapładniają się wzajemnie ideami i pomysłami. Kreatywność ogromnie wzrasta, tempo innowacji ulega wielkiemu przyspieszeniu. Kluczowa kategoria ekonomii, jaką jest cena produktu – podstawa konkurencyjności, ustępuje miejsca innowacjom jako najważniejszemu jej kryterium. Społeczeństwo komunikacyjne jawi się jako *bonne société*

Wspomniany Arun Netravali Szef Laboratoriów Bella, prognozuje, że pokolenie dziś przychodzące na świat rozpocznie dorosłe życie w rzeczywistości, w której inteligentne sieci oplotą planetę niczym skóra. Czujniki rozmieszczone wszędzie będą przekazywać wszelkie informacje wprost do sieci, jak nerwy transmitujące informacje do mózgu. Ziemia będzie wtedy inną, lepszą planetą.

Do grona orędowników Sieci dołączył pionier badań nad społecznościami wirtualnymi, Howard Rheingold. W swej ostatniej książce („Inteligentne masy”, 2002) manifestuje swą wiarę w to, że społeczeństwo masowe przyszłości będzie mobilne i twórcze. Poglądy Rheingolda są godne odnotowania, ponieważ stoją w opozycji do powszechnego przekonania badaczy, że technologie informacyjne kładą kres społeczeństwu masowemu. To społeczeństwo, które narodziło się epoce przemysłowej, formowało jednostkę jako typowy egzemplarz gatunku wskutek bombardowania wielomilionowych audytoriów tymi samymi przekazami przez wielkie mołochy medialne. Wierzone, że interaktywne technologie niszczą masę i tworzą społeczeństwo pomasowe, którego zawołaniem jest *customization* – obsłużenie każdego wedle jego upodobań. Otóż, zdaniem Rheingolda tak się nie stanie, społeczeństwo sieciowe będzie masowe, ale będzie to inteligentne społeczeństwo masowe, dzięki kolektywnej inteligencji roju (*swarm*

intelligence). Wezwania „przyłącz się, włącz się, dołącz” stają się słowami kluczowymi tego nowego paradygmatu.

Ze wszystkich tych opisów wyłania się budujący obraz. Jest to wizja *communicative society*, którą za Habrmasem rozwijają inni badacze (Stevenson, June 1995 5-36). Jeśli w XXI w. ma się ono ziścić, to musi przyjąć inny światopogląd niż dominujący w epoce maszyn: holistyczny, twórczy, przyjazny, demokratyczny, skoncentrowany na człowieku, partnerski. Zakłada to akceptację różnych rzeczywistości, w których, sensy i znaczenia będą społecznie konstruowane, a nie naturalnie immanentne dla świata. Z dorobku epok zachowamy orientację na indywidualizm, ale indywidualizm społeczny, jak sobie go wymarzył Immanuel Kant. Również technologia będzie konstruktem społecznym ucieleśniającym wartości, przekonania i ideologie użytkowników przez uczestniczące projektowanie i wartościowanie

System społeczno-polityczny zachowa pożądaną stan *harmonia mundi*: równowagę między człowiekiem, ludźmi, naturą i technologią. Rozwinie się demokracja partycypacyjna dzięki mechanizmom podejmowania decyzji na poziomie społeczności lokalnych. ITs pomogą w uprawnianiu ludzi. Państwo narodowe będzie coraz bardziej uzupełniane przez organizacje obywatelskie, pozarządowe i z czasem przez nie zastępowane..

System społeczno-ekonomiczny musi się sam podtrzymywać (rozwój wspólnotowy i lokalne kooperatywy oparte na samorządzeniu i wymianie informacji/wiedzy). Powstaną ekoregiony rozwijające rynek prosumencki (producent i konsument „w jednym”, co zapowiadał już Alvin Toffler w „Trzeciej Fali”), co wymagać będzie m.in. redefinicji pracy. Poziom sprawiedliwości społecznej i równości będzie wykorzystywany do pomiaru wskaźników jakości życia. Ruchy feministyczne zrestrukturyzują relacje między płciami; wpłyną na to, że technologie będą służyć potrzebom człowieka, a nie człowiek będzie się musiał do nich przystosować.

W kulturze nastąpi wzrost różnorodności. Zagrożone kultury przetrwają i powstaną nowe. Będzie to nie tyle globalna wioska, co „wioski na globie”. Technologia ułatwi komunikację międzykulturową. Cenione będą wiedza i twórczość. Współpraca wyprze konflikty i przemoc. Najbardziej usieciowione społeczeństwa będą zorientowane na wartości postmaterialne – intelektualne i duchowe, jakość życia.

W systemach organizacyjnych przedsiębiorstw ludzie będą najważniejszym czynnikiem, ale nie jako zasoby, które trzeba wyeksploatować, po czym jak zmurszałą cegłę w murze wymienić na inne. Systemy będą nastawione na twórcze nauczanie promujące tworzenie wiedzy i dzielenie się nią. Edukacja będzie skoncentrowana na uczniu, partycypacyjna, ustawiczna, zachęcająca do interaktywności. Uczenie się będzie cenione dla samego uczenia, a nauczyciele mają to ułatwić z pomocą interaktywnej edukacji oraz treningu *in situ* i w domu; technologia edukacyjna zaspokoi zróżnicowane potrzeby użytkownika.

Interaktywne i twórcze praktyki kulturalne wypełnią czas wolny; będzie większy wybór i różnorodność programów kierowanych do węższych grup, a współtworzonych także przez odbiorców. Rozrywka przywróci nam poczucie człowieczeństwa i pozwoli na rekonceptualizację świata w duchu naszych potrzeb.

W myśleniu o przyrodzie zwycięży podejście koewolucyjne (jako gatunek ewoluujemy wspólnie z innymi, nie wynosimy się ponad nie), samopodtrzymująca się eksploatacja zasobów, zapobieganie skażeniom środowiska, a nie tylko łagodzenie ich skutków, jak to ma miejsce dziś. Nastąpi rozwój ekotechnologii i recylingu, rolnictwa nisko-kosztowego, powstaną bioregiony cieszące się dużą autonomią. Działania ludzkie wchodzą w synergiczny układ z procesami zachodzącymi w ekosystemach.

W Refleksji nad przyszłością będzie się pod uwagę alternatywne przyszłości; najważniejsza będzie wspomniana koewolucja oraz samoorganizacja i samoregulacja, system samoodnawialny oparty na synergii nie degradujących natury i kultury jako czynników rozwoju.

Taka wizja spodobałaby się z pewnością wszystkim zwolennikom globalnego społeczeństwa obywatelskiego, które miałoby kontrolę nad obydwoma siłami mu zagrażającymi: wielkim biznesem i strukturami władzy. Ekspansji obu tych sił nie udaje się jak dotychczas powstrzymać. Spodobałaby się wszelkim ruchom alternatywnym: feministycznym, antyglobalnym, ekologicznym i in. Ale czy jest realna?

W wielu opisach współczesnego świata, w których roi się od horrorów, pojawiają się jednak jaskółki nowego. Ryszard Kapuściński, który ma ucho na te nowinki powiada (2002: 19) „Świat tak szybko się zmienia, że nasza świadomość za nim nie nadąża. A tu trzeba rewolucji wyobraźni! /.../ Nowość, którą możemy obserwować, to narodziny w skali świata nowych instytucji i nowych form społecznej ekspresji. Przeżywamy kryzys tradycyjnych partii politycznych i związków zawodowych. One istnieją często tylko na papierze. Społeczeństwa się zmieniły, są dziś wielozawodowe, wielokulturowe, wielowszystko, do tego stopnia, że tradycyjne formy zrzeszania się, zabiegania o swoje prawa nie odpowiadają już na zapotrzebowania ludzi, na ich dążenia. Rodzą się nowe struktury, nowe formy organizacji”. Dużą rolę w powstawaniu tych nowych form odgrywają sieci społeczne, które zyskały potężne wsparcie w ITs.

Czy to nie jest utopia. Pierre Levy odpowiada: nawet jeśli tak, to jest to utopia osiągalna - *une utopie realisable*. Wybitny filozof polityki, Karl Popper, przestrzegał, że ten, kto stworzyć niebo na ziemi, sprowadza raczej piekło. Ale George Bernard Shaw był innego zdania; powiadał, że utopie są jak gwiazdy, nawet jeśli nie można ich dotknąć, to można i trzeba się na nie kierować. Historia zachodniej cywilizacji obfituje w dobre i złe utopie. Te dobre odegrały pozytywną rolę, uwarżliwiały na ważne dla ludzkości problemy, wyznaczały punkty orientacyjne; te złe były przestrogą dla następnych pokoleń, aby nie eksperymentować na ludziach.

W przypadku sieci nikt nie chce na nikim eksperymentować. Ich powstawanie jest procesem ciągle jeszcze bardziej spontanicznym niż organizowanym. Jeśli z takiego spontanicznego procesu zrodziłoby się coś dobrego, to tym lepiej dla ludzi. Kooperatywne sieci są najlepszą receptą na rosnącą złożoność świata, którą można skuteczniej zarządzać raczej przez samoregulację niż regulację z zewnątrz.

Optymiści idealizują sieci w przekonaniu, że zepchną one w cień reżimy regulacji oparte na hierarchii. Nawet jeśli się nie podziela tego optymizmu, to można go zrozumieć – jest w nim duży ładunek tęsknoty za jakimś dobrym ładem. Ludzie łakną kooperacji w świecie wszechogarniającej konkurencji. Wszystko pięknie. Ale optymizm nie powinien się kłócić z realizmem. W poważnych scenariuszach społeczeństwa sieciowego (m.in. u M. Castellsa) jest miejsce na konflikty, ale najczęściej spycha się je na drugi plan jako nie bardzo pasujące do paradygmatu. Nawet jeśli się będą one pojawiać, to kooperacyjno-sieciowy system regulacji się z nimi upora. Nie będzie natomiast miejsca na wyniszczające wojny sieciowe. Otóż to jest najsłabsza strona tego scenariusza. Bo jeśli nawet nie będzie takich konfliktów w sieciach cywilizacji zachodu, to będzie zagrożenie ze strony „netdżihadu”. Z takim transferem *infowars* do cyberprzestrzeni mamy však już dziś do czynienia.

3. SPOŁECZEŃSTWO KONTYNUACJI – DEJA VU, CZYLI OSWAJANIE NOWYCH TECHNOLOGII W DOBRYM I ZŁYM

Argumentacja rzeczników z tego nurtu przedstawia się z grubsza tak: ludzkość przechodziła już wiele rewolucji technicznych i wszystkie zostały jakoś oswojone. Oczywiście, nasze życie się zmienia i będzie się zmieniać, tak jak zmieniał się pod wpływem druku, telefonu, automobila, czy aeroplanu, ale te wynalazki nie zdominowały człowieczeństwa, nie zerwały ciągłości kultury, bywały skoki, ale w sumie zwyciężała ewolucja. Będzie zatem więcej kontynuacji niż myślimy. Człowiek pozostanie człowiekiem ze swymi dobrymi i złymi cechami, nie da się przerobić, ani w diabła ani w anioła. Dotyczy to również systemów społecznych, którego będą jak zawsze kombinacją *desirable* i *undesirable*.

Często zdarza się i tak, że po wielkiej zmianie technologicznej następuje pewna dwoistość: powstają jakby dwa społeczeństwa w jednym: beneficjanci absorbują zmiany, gruntują je; ci którzy nie potrafią się do nich zaadaptować – odrzucają. Interpenetracja idei, stylów życia, pracy, zarządzania itp. między obu układami: starym i nowym jest raczej nikła, a efekty dyfuzyjne nieznaczne. Funkcjonują dwa światy wartości wzorów i instytucji.

Słowem: nowe technologie mogą utwierdzać stary, typowy dla poprzednich etapów cywilizacji system społeczno-ekonomiczny i kulturowy. Wcale nie muszą go burzyć.

Wynikałoby z tego, że do kategorii mrzonek należałoby zaliczyć na przykład wizje dehierarchizacji struktur społecznych i ekonomicznych, politycznych, kulturowych, czyli „gatesism” (od Billa Gatesa) zamiast fordyzmu i taylorizmu (taśma produkcyjna). Widzi się w tym wszelkie znamiona rewolucji organizacyjnej. Wielkie korporacje zmieniają swą strukturę, stają się bardziej zdecentralizowane. Pojawił się fenomen przedsiębiorstw wirtualnych, które funkcjonują jedynie w Internecie, co sprawia, że najistotniejsza i przyszłościowa część gospodarki staje się sieciowa. Taka skomputeryzowana organizacja kieruje się własną logiką

Mówi się o tych przedsiębiorstwach, że są one ze swej natury informacyjne o symbolicznej infrastrukturze. Informatyzacja procesów wytwórczych miałaby oznaczać zupełnie inną niż w społeczeństwie przemysłowym filozofię zarządzania inteligentnymi firmami nazywanymi *intelprises* i operującymi na „rynku intelektualnym” (*intelplace*).

Mowa także o tym, że mamy do czynienia z decentralizacją decyzji z centrum ku lokalnym ośrodkom produkcji i usług, co umożliwia technologia zarządzania w sieci. Obfitość kanałów komunikacyjnych, wzrost konkurencyjności, multiplikacja rynków, deregulacja etc. oznacza, przesuwanie „władzy na skraj organizacji”. Sieć zmniejsza znaczenie ogniw pośrednich, których zadaniem było przekazywanie informacji z góry w dół systemu. Konkurencja, zwiększanie produktywności w świecie takiej ekonomii się. Mobilność informatycznych środków produkcji osiągnęła niespotykaną wcześniej skalę.

Tymczasem można sobie doskonale wyobrazić społeczeństwo informacyjne jako nadbudowę nad „taśmą elektroniczną” i formy kontroli nad pracą umysłową; musi się pracować w odpowiednim algorytmie; jeśli ktoś ma zachciankę robić coś innego przy komputerze to system od razu to ujawni. Komputer wymaga dyscypliny, żeby można było zarządzać „przepływami pracy”. Jeśli ktoś tego nie rozumie, to zachowuje się jak dezenter porzucający taśmę przemysłową. Różnice między jedną taśmą a drugą nie wydają się jakościowe, najważniejsza jest taka, że przy jednej operuje się narzędziami mechanicznymi, przy drugiej zaś symbolami, co wymaga bardziej skomplikowanych czynności (choć dla tych, którzy to potrafią, to już rutyna), ale narzędzia ma się przy sobie –

w pierwszym przypadku w licznych kieszonkach uniformu, w drugim na twardym dysku lub w sieci.

Niezapomniany Charlie Chaplin ośmieszył taśmę w filmie „Nowe czasy”, pokazał nędzną kondycję człowieka „zfordyzowanego” – wyalienowanego, zredukowanego do śrubki w mechanizmie, prostej funkcji. Nie spotkałem się jeszcze z wybitnym filmem, który by ośmieszał *homo informaticus*, choć coraz więcej żartów i satyry powstaje na ten temat.

Pojawienie się przełomowej technologii wyzwało na początku na ogół skrajne emocje. Tak było w przypadku większości wynalazków epoki przemysłowej. Po jakimś czasie jednak okazuje się, że siła inercji była większa, niż się wydawało. Dotyczyło to nie tylko wynalazków, ale także idei. Rewolucjoniści zawsze święcie wierzyli, że odmieniają świat, radykalnie zrywają z przeszłością. Nawet jeśli wygrywali, to często sami padali ofiarą: rewolucja zjadała swe dzieci. W następnym pokoleniu na ogół wytracał się impet rewolucyjny, stara kultura odżywała jak przydeptana trawa. Tak było w przypadku restauracji wielu instytucji feudalizmu w porewolucyjnej Francji, podobnie jak w przypadku powrotu samodzierżawia w bolszewickiej Rosji.

4. ZAKOŃCZENIE

Historia zmian społecznych, do których dochodzi najczęściej pod wpływem nowych narzędzi, ukazuje pewien uniwersalny wzór: cztery możliwe reakcje kultury na innowację: **akceptację, odrzucenie, hybrydyzację i dualizm**. Pierwsze dwie, opisywane najczęściej, występują jednak rzadko w czystej postaci. Na ogół to, co następuje po zmianie, jest kombinacją dwóch ostatnich scenariuszy. Hybrydyzacja to fuzja starego z nowym: społeczeństwo wychodzi z punktu A, ale nie dochodzi do punktu B, czyli nowego stanu. Zatrzymuje się gdzieś po drodze wykształcając pewien pośredni, właśnie hybrydyczny model. Często zdarza się i tak, że po wielkiej zmianie technologicznej następuje pewna dwoistość: powstają jakby dwa społeczeństwa w jednym: beneficjanci absorbują zmiany, gruntują je, ci którzy nie potrafią się do nich zaadaptować – odrzucają. Interpenetracja idei, stylów życia, pracy, zarządzania itp. między obu układami: starym i nowym jest raczej nikła, a efekty dyfuzyjne nieznaczne. Funkcjonują dwa światy wartości wzorów i instytucji. Głośny przed laty autor książki „Małe jest piękne”, Ernst Freidrich Schumacher przytaczał mnóstwo przykładów ze świata zacofanego, gdzie budowano nowoczesne, zautomatyzowane fabryki, a obok żyły sobie po staremu całe społeczności plemienne, na które te fabryki nie miały żadnego wpływu. W ten sposób objawiała się siła długiego trwania.

Pozostawiam osądowi Państwa, który z tych scenariuszy ma największe szanse spełnienia, ziarna jakiego świata kiełkują już w naszej rzeczywistości. Można zaryzykować stwierdzenie, że żaden z nich nie zrealizuje się w czystej postaci i świat będzie płataniną wszystkich tych trzech tendencji.

BIBLIOGRAFIA

- [1] Brzeziński Z., Kłopoty dobrego hegemonu, „Gazeta Wyborcza”, 4-5.07.1998
- [2] Chyła W. [2002], Człowiek w dobie mediów, media w dobie technologii (maszynopis)
- [3] Kapiuściński R., w wywiadzie dla „Gazety Wyborczej „Ciszej z tymi werblami”: 24-26.12 2002, s. 19):
- [4] De Kerkhove D., Powłoka kultury, Zysk I S-ka, Poznań 1999
- [5] Levy P. [1996]: The Second Flood. Report on Cyberculture, Council of Europe, Paris,

- [6] Postman N. Technopol. Triumf techniki nad kulturą", Warszawa, 1995,
- [7] Rheingold H., Smart Mobs. The Next Social Revolution. Transforming Cultures and Communities in the Age of Instant Access, Basic Books, Cambridge Mass, 2002,
- [8] Talbot C. [1995], The Future does not compute – Transcending the Machines in Our Midst, Sebastopol, Ca,
- [9] Stevenson T., June L., *Conventional, Communicative and Artificial Age*, Future Research Quarterly", Fall 1995: 5-36

GLOBAL INFORMATION SOCIETY: THREE MEGATRENDS

Summary

The three megatrends analyzed in the paper are as follows:

1. Matrix, or monitored space

Humans will not be any more a reliable source of information for the systems collecting data on them, to make them even more algorithmic, as the information and communication technology itself is seen more reliable and creditworthy. This lack of confidence in humans accounts for the practice of engaging computers to obtain trustworthy data, or when to control and measure the productivity, integrity and probity, and commitment of an employed person. It is becoming more profitable to derive information from objective systems independent from individuals. There are however some people who hold strong convictions that this will deepen the erosion of trust between people, thus weakening social capital.

2. Bonne Societe, or desirable social order

The culture and ideas are increasingly created in networks and spread out within these. These networks are reinforced by information and communication technologies as those of freedom; this is the text message generation. Cultural activities flow out of institutions into networks in which the social and cultural capital is vented; networks of self-organising and self-regulating people, creating cultures. These cultures show a high degree of individuality and specific communication codes, which make enormously for cultural variety. Citizenship is defined increasingly by cultural features, affiliation to *choice groups* rather than to nation-states. Information and communication technologies allow people to return to spontaneous communication. Billions of text messages, and the more recent multimedia equivalents, emails, chats and blogs (online diaries) and dozens of thousands of discussion groups on the internet, lead to the belief that in the age of individualism, a certain form of *collective intelligence and culture* is emerging.

3. Déjà vu or more continuity than change

The information society will not really be smarter than the industrial one. The conventional wisdom about the information society is that it will enhance freedom and individualism. The reality seems different; the information society will establish its own electronic pecking order. It will be heading for a superstructure over the electronic assembly line and control over intellectual work. If anyone has a whim to do something else, the system will detect it immediately. Computers demand algorithmic discipline in the teamwork to manage and control the *work flows*. If anyone does not understand this logic they behave like a deserter abandoning the assembly line. The difference between the assembly line of the industrial era and that of the information age is not qualitative. The former was based on workers using mechanical tools, and at the latter will be based on symbolic tools that demand more complex activities. In both cases the tools are at hand, in the first case hand tools in pockets, in the second, on a hard disc or on the internet.

Ilona Bluemke, Adrian Kuszewski, Konrad Łapienko

Instytut Informatyki, Politechnika Warszawska

KONCEPCJA SYSTEMU ZDALNEGO NAUCZANIA W JAVIE

Streszczenie

W pracy przedstawiono koncepcję systemu zdalnego nauczania zrealizowanego w Instytucie Informatyki PW. System ten zaprojektowano korzystając z narzędzia Rose i zrealizowano w języku Java, co daje niezależność od platformy sprzętowo – programowej. Zastosowano architekturę typu klient-serwer. Wykorzystano mechanizm zdalnego wywołania procedur Java RMI (*Remote Method Invocation*) rozwiązujący problemy komunikacji w sieci. Ze względu na łatwość realizacji zdecydowano się na realizację systemu asynchronicznego. Wymiana informacji między nauczycielem a studentami nie odbywa się w czasie rzeczywistym. System umożliwia użytkownikom komunikację między sobą poprzez moduł wewnętrznej poczty elektronicznej. System łączy funkcjonalność poczty elektronicznej oraz FTP. Istotną przewagą systemu jest jednak ścisła kontrola uprawnień użytkowników przy podejmowaniu przez nich jakichkolwiek akcji.

1. WSTĘP

Uczenie na odległość (ang. *distance learning*) zostało wynalezione już prawie 100 lat temu. Prawdziwy przełom nastąpił jednak wraz z upowszechnieniem się internetu. To nowe medium dało nauczaniu na dystans nowe, nie spotykane dotąd możliwości interakcji i komunikacji uczącej się grupy oraz bliski nauczaniu tradycyjnemu, rzeczywisty kontakt studentów z prowadzącym kurs nauczycielem.

W ostatnich latach powstaje coraz więcej systemów zdalnego nauczania. Realizacja systemów zdalnego nauczania może odbywać się w oparciu o specjalizowane, zintegrowane pakiety np. Lotus Learning Space, WebCT. Przegląd realizacji systemów zdalnego nauczania zawiera praca [1]. Zalety tego typu systemów to wielofunkcyjność np. możliwości tworzenia kursów i zarządzania kontami użytkowników, zaawansowane możliwości zdalnej współpracy, oraz możliwość użycia przez osoby nie posiadające wiedzy z zakresu programowania. Systemy te są jednak bardzo drogie. Systemy zdalnego nauczania mogą być realizowane także w postaci stron WWW ze wsparciem JavaScript lub appletów w języku Java. Zaletami takich systemów są niezależność od platformy sprzętowo-programowej oraz niski koszt, wadą jest natomiast ograniczona funkcjonalność systemu. Inne rozwiązania mogą mieć postać kursów multimedialnych. Jeszcze innym rozwiązaniem może być realizacja systemu zdalnego nauczania w języku programowania wysokiego

poziomu np. C++, Delphi, czy Java. Takie właśnie podejście zastosowano realizując system zdalnego nauczania w Instytucie Informatyki PW. System zdalnego nauczania zaprojektowano korzystając z narzędzia Rose. Jest on zrealizowany w języku Java, co daje niezależność od platformy sprzętowo – programowej. Zastosowano architekturę typu klient-serwer. Wykorzystano mechanizm zdalnego wywołania procedur Java RMI (*Remote Method Invocation*) rozwiązujący problemy komunikacji w sieci. Ze względu na łatwość realizacji zdecydowano się na realizację systemu asynchronicznego. Wymiana informacji między nauczycielem a studentami nie odbywa się w czasie rzeczywistym. Asynchroniczność pociąga za sobą brak kontaktu studenta z nauczycielem oraz współuczniami, którzy mogli by służyć radą i poprawiać błędy. W systemie wyodrębniono dwa typy użytkowników tj.: studentów, którzy mogą zostać zorganizowani w ramach tzw. grup studenckich i wykładowców. Każda grupa składa się z dokładnie jednego wykładowcy, prowadzącego grupę i dowolnej liczby studentów. Materiały dydaktyczne są zorganizowane w katalogi, których właścicielami mogą być jedynie wykładowcy. Udostępnianie materiałów dydaktycznych z danego katalogu odbywa się poprzez określenie przez jego właściciela grup studenckich, których członkowie będą uprawnieni do ich pobierania. System umożliwia użytkownikom komunikację między sobą poprzez moduł wewnętrznej poczty elektronicznej. System łączy więc funkcjonalność poczty elektronicznej oraz FTP. Istotną przewagą systemu jest jednak ścisła kontrola uprawnień użytkowników przy podejmowaniu przez nich jakichkolwiek akcji. W przypadku wewnętrznego modułu poczty elektronicznej przyjęto założenie, że komunikacja odbywa się wewnątrz grupy. Eliminuje to problem zaśmiecania skrzynki przez innych użytkowników systemu.

Nadzór wykładowcy nad procesem nauczania studentów jest realizowany poprzez rejestrację aktywności w systemie wszystkich jego użytkowników. Każdy z wykładowców ma później prawo prześledzić rejestr aktywności studentów, którzy mu podlegają.

W rozdziale drugim zawarto wybrane informacje o realizacji systemu. Mechanizm RMI skrótowo przedstawiono w rozdziale trzecim a podstawowe funkcje systemu pokazano w rozdziale czwartym.

2. IMPLEMENTACJA SYSTEMU ZDALNEGO NAUCZANIA

Jednym z podstawowych założeń implementacyjnych systemu zdalnego nauczania była niezależność od platformy sprzętowo-programowej. Chodziło także o to, by system umożliwiał implementację założonej funkcjonalności systemu, przedstawionej w rozdziale czwartym. Powyższe wymagania mogą zostać spełnione przy wykorzystaniu jako języka implementacji języka Java. Dodatkowymi zaletami takiego rozwiązania są: relatywnie niski koszt realizacji systemu i wsparcie ze strony narzędzi CASE w procesie jego projektowania. Projekt systemu zrealizowano w narzędziu Rose firmy Rational, co znacznie przyspieszyło projektowanie i realizację systemu, a autorom pozwoliło na zdobycie doświadczenia w korzystaniu z narzędzi CASE.

Istotnym atutem języka Java, przy realizacji systemów pracujących w środowisku rozproszonym, jest również bogaty zestaw oferowanych mechanizmów komunikacji sieciowej (gniazda, RMI, CORBA). Ze względu na asynchroniczny tryb pracy implementowanego systemu zdalnego nauczania zdecydowano się na zastosowanie mechanizmu zdalnego wywołania procedur (RPC), którego odpowiednikiem w języku Java jest RMI (*Remote Method Invocation*), opisany w rozdziale 3.

Zrealizowany system korzysta z następujących danych:

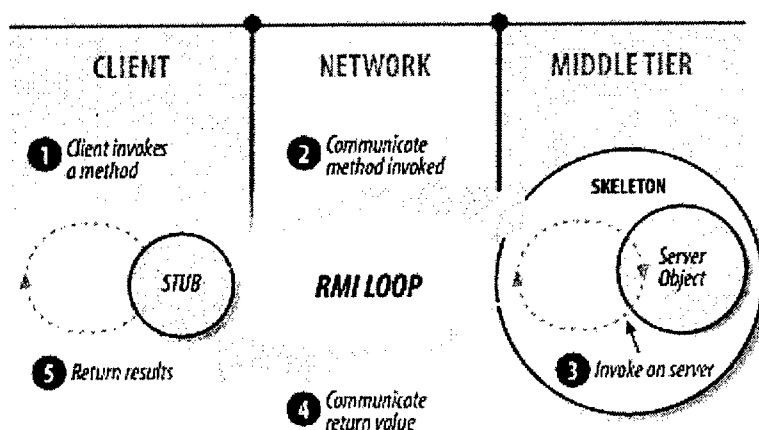
- Dane serwera – Plik konfiguracyjny zawierający informacje o serwerze (adres IP i numer portu), ścieżki do katalogów z kontami i materiałami dydaktycznymi.
- Baza kont – Każde konto jest reprezentowane jako jeden katalog, w którym znajdują się pliki przechowujące informację o zawartości skrzynki pocztowej, logach i o użytkowniku (login, hasło, dane osobowe i jego przynależność do grup studenckich). W osobnych plikach (poza strukturą katalogową kont) są przechowywane informacje o wszystkich kontach zarejestrowanych w systemie oraz wszystkich grupach studenckich (identyfikator i opis grupy, login wykładowcy prowadzącego grupę, lista identyfikatorów studentów do niej przypisanych).
- Baza materiałów dydaktycznych – Fizyczna struktura katalogów odzwierciedla strukturę katalogów z materiałami dydaktycznymi udostępnianymi w systemie. Każda lekcja w systemie jest reprezentowana przez pojedynczy plik umieszczony w odpowiednim katalogu. Każdy katalog w systemie jest scharakteryzowany przez identyfikator, opis, identyfikator właściciela. Ponadto z każdym z nich związane są trzy listy: lista lekcji w katalogu, lista wykładowców o autoryzowanym dostępie do katalogu oraz lista grup studenckich mających dostęp do katalogu. Te informacje są przechowywane w bliźniaczej strukturze katalogów, co stanowi dodatkowe zabezpieczenie przed ich utratą.

3. MECHANIZM RMI

Mechanizm RMI (*Remote Method Invocation*) [2] jest jednym z mechanizmów języka Java służących do realizacji komunikacji sieciowej. Technologia ta pozwala korzystać z obiektu znajdującego się na zdalnym komputerze w ten sam sposób, w jaki korzysta się z obiektów lokalnych. Kod związany z przesyłaniem komunikatów przez sieć jest generowany automatycznie przez odpowiedni kompilator (*rmic*). Znakomicie upraszcza to realizację rozproszonych systemów asynchronicznych zwalniając programistę z konieczności zajmowania się problemami komunikacji sieciowej. Pozwala to na zdecydowanie łatwiejszą realizację systemu o architekturze klient-serwer w stosunku do tradycyjnego, synchronicznego mechanizmu gniazd.

Po skompilowaniu pliku klasy obiektów zdalnych z otrzymanego pliku *.class* należy wygenerować, przy pomocy kompilatora *rmic*, pliki namiastki (*stub*) i szkieletu (*skel*) odpowiadające za komunikację sieciową. Niezbędny do poprawnego działania mechanizmu RMI plik namiastki (*stub*) powinien być dostępny (poprzez protokół HTTP lub FTP) dla programu – klienta. Umożliwia to automatyczne jego pobranie przy każdej modyfikacji strony serwera. Udostępnienie jest możliwe przez podanie przy uruchomieniu serwera, jako parametru maszyny wirtualnej (*-Djava.rmi.server.codebase*), adresu, pod którym został umieszczony dany plik. W podobny sposób można udostępnić inne pliki wymagane do poprawnej pracy systemu. Ze względu na mechanizm ochrony danych dla aplikacji rozproszonych w języku Java, jest konieczne określenie poziomu ich dostępu do zasobów na maszynie lokalnej. Możliwe jest to poprzez wyspecyfikowanie tzw. polityki bezpieczeństwa w postaci pliku uprawnień (*policy file*), a umiejscowienie pliku jest podawane przy uruchomieniu aplikacji jako parametr maszyny wirtualnej (*-Djava.security.policy*).

Udostępnienie obiektu zdalnego dla klientów wymaga jego rejestracji w rejestrze RMI, specyfikuje się przy tym adres maszyny serwera, numer portu oraz referencję do zdalnego obiektu. Tak zarejestrowany obiekt jest od tej chwili dostępny, poprzez jeden z implementowanych przez niego zdalnych interfejsów. Ideę działania mechanizmu RMI pokazano na rys. 1 zaczerpniętym z [2].



Rys. 1. Mechanizm działania mechanizmu RMI (źródło: [2])

4. FUNKCJONALNOŚĆ SYSTEMU

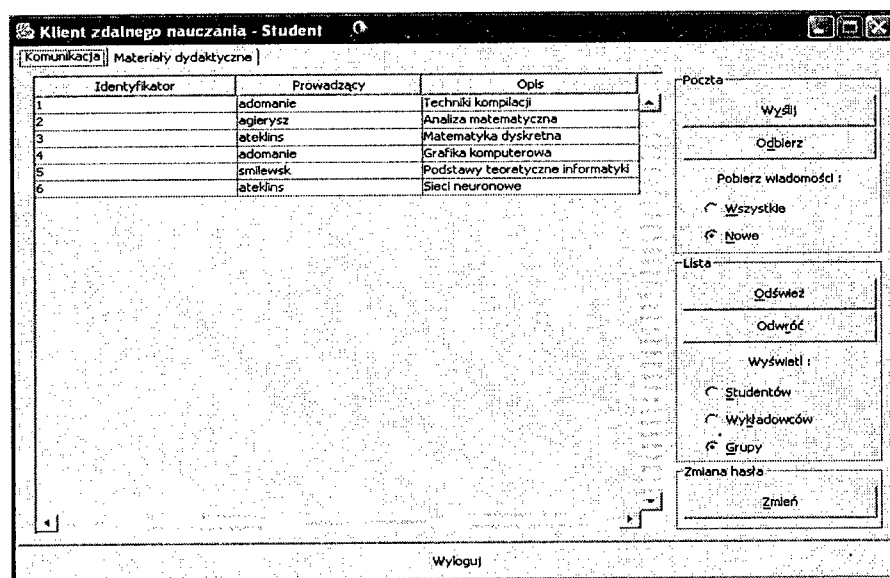
System zdalnego nauczania został zrealizowany w architekturze klient-serwer. Możliwości systemu zostaną przedstawione osobno dla każdej części. W rozdziale 4.1 przedstawiono główne funkcje strony klienta, a w rozdziale 4.2 strony serwera. W systemie są trzy kategorie użytkowników: studenci, wykładowcy i administrator. Usługi systemu są przedstawione osobno dla każdej kategorii.

4.1. Funkcje klienta

Po stronie klienta wspólnymi usługami dostępnymi dla studentów i wykładowców są logowanie, wylogowanie, zmiana hasła, zmiana adresu i portu serwera. Kolejna bardzo istotna funkcja to asynchroniczna komunikacja z użytkownikami pozwalająca na:

- przysyłanie wiadomości do skrzynek pocztowych poszczególnych użytkowników i grup studenckich (w tym możliwość wysyłania jednego załącznika),
- znakowanie wiadomości – przeczytane/nieprzeczytane,
- odbieranie/usuwanie wiadomości.

Na rysunku 2 pokazano okno komunikacji studenta, (podobnie wygląda okno wykładowcy). Pozostałe funkcje są różne dla wykładowców i studentów i zostały przedstawione w rozdziałach odpowiednio 4.1.1 i 4.1.2.



Rys. 2. Panel komunikacji – interfejs studenta

4.1.1. Usługi dla wykładowców

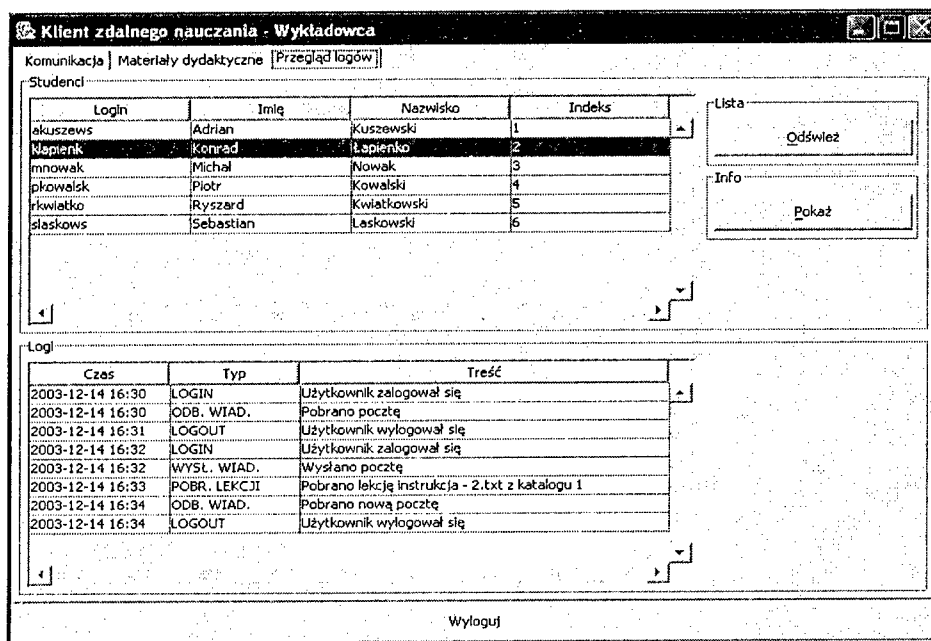
Jedną z podstawowych funkcji dostępnych dla wykładowców jest zarządzanie materiałami dydaktycznymi. W zakres tej funkcji wchodzi:

- tworzenie/usuwanie katalogów z materiałami dydaktycznymi,
- dodawanie/usuwanie lekcji do/z własnych katalogów,
- zmiana praw dostępu do własnych katalogów dla wybranych wykładowców,
- zmiana praw dostępu do katalogów dla wybranych, prowadzonych przez wykładowcę, grup studenckich.

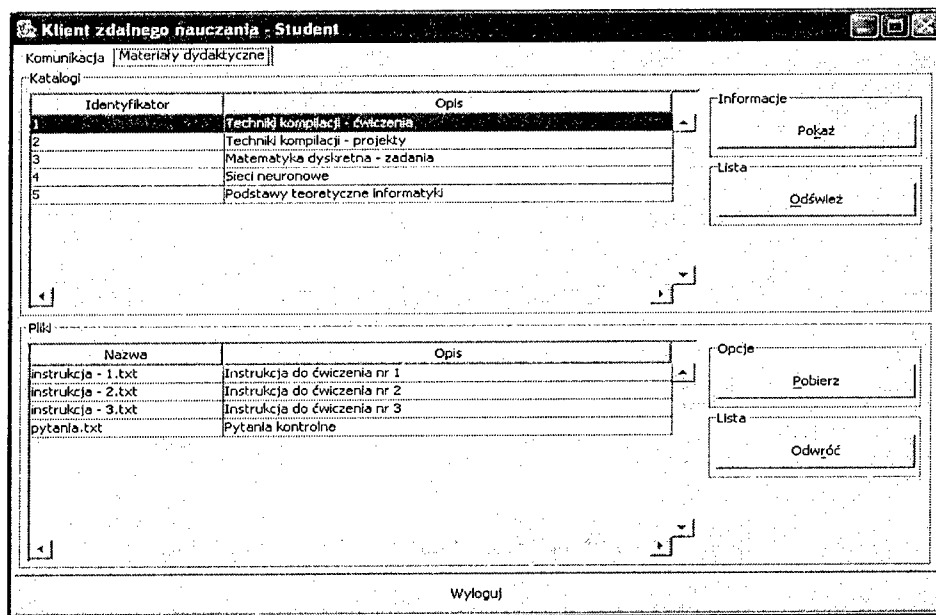
Inne funkcje wykładowcy to dostęp do informacji o studentach należących do prowadzonych przez wykładowcę grup, do informacji o wszystkich wykładowcach. Poprzez przeglądanie logów studentów (rys.3) wykładowca ma możliwość nadzorowania, czy raczej śledzenia procesu nauczania.

4.1.2. Usługi dla studentów

Jedną z podstawowych funkcji dostępnych dla studentów jest pobieranie materiałów dydaktycznych czyli pobieranie lekcji z udostępnionych (poprzez grupy) katalogów, a także dostęp do informacji o studentach należących do tych samych grup, co dany użytkownik. Student ma także dostęp do informacji o wykładowcach prowadzących grupy, do których on należy. Na rysunku 4 pokazano przykładowe okno materiałów dydaktycznych.



Rys. 3. Panel nadzoru procesu nauczania – interfejs wykładowcy



Rys. 4. Panel materiałów dydaktycznych – interfejs studenta

4.2. Funkcje serwera

Po stronie serwera jest możliwość nadzoru i rejestracji wszystkich operacji systemu. Strona serwera oferuje usługi dla administratora. Administrator ma, tak jak każdy użytkownik systemu, możliwość komunikowania się z innymi użytkownikami poprzez wewnętrzną pocztę systemu. Główne funkcje administratora dotyczą zarządzania kontami użytkowników i serwerem. Zarządzanie kontami polega na dodawaniu i usuwaniu kont, zmianie haseł oraz danych użytkowników, a także na dodawaniu, usuwaniu i na modyfikacji przydziału studentów do grup. Na rysunku 5 pokazano okno dodawania grupy studenckiej, a na rysunku 6 okno zarządzania kontami. Zarządzanie serwerem sprowadza się do jego uruchamiania lub zatrzymania (co jest równoznaczne z zakończeniem działania programu) i na zmianie numeru portu, na którym serwer oczekuje na połączenia.

Dodaj nową grupę studencką

Wybierz wykładowcę który poprowadzi grupę

Login	Imię	Nazwisko
anowak	Anna	Nowak
kmronka	Katarzyna	Monka

Podaj identyfikator nowej grupy (musi być to liczba)

12

Podaj krótki opis grupy

Grupa nr 12

Dodaj grupę

Zatwierdź zmiany

Anuluj zmiany

dane nowej grupy (aby zmiany zostały zapisane do bazy naciśnij Zatwierdź zmiany)

Rys. 5 Zakładanie nowej grupy studenckiej

Zarządzanie grupami studenckimi i kontami użytkowników

Zarządzanie kontami | **Zarządzanie grupami** | **Komunikacja z użytkownikami**

Login	Imię	Nazwisko	Nr indeksu
abek	Adam	Bek	100005
akowalsk	Artur	Kowalski	100001
izajac	Joanna	Zajac	100004
kwozniak	Krzysztof	Wozniak	100002
wkonar	Władysław	Konarski	100006
zpawlak	Zenon	Pawlak	100003

Wyświetl listę:

☒ **Studentów**

☐ **Wykładowców**

Odśwież

Konta:

Dodaj nowe

Usuń

Konto:

Zmień hasło

Zmień dane

Zmodyfikuj dane kont (wybranego z listy)

Przeładuj logi

Wyczyść logi

Zapisz stan serwera

Zakończ

Rys.6. Zakładka „Zarządzanie kontami” – zmiana informacji o użytkownikach

5. ZAKOŃCZENIE

W pracy przedstawiono koncepcję systemu zdalnego nauczania zrealizowanego w Instytucie Informatyki PW. Jego podstawowym zadaniem jest wspieranie procesu nauczania w środowisku akademickim. Z tego względu sposób zorganizowania użytkowników w systemie odwzorowuje strukturę grup studenckich znaną z większości uczelni. Takie założenie może jednak ograniczać jego zastosowanie w innych środowiskach. Ze względu na realizację w języku Java system może działać na różnych platformach. Dzięki zastosowaniu ogólnego podejścia do zagadnień komunikacji i współdzielenia materiałów dydaktycznych, uzyskano uniwersalność systemu pozwalającą na jego wykorzystanie do dowolnego tematu procesu kształcenia. Założenia systemu obejmowały realizację tylko pewnej podstawowej funkcjonalności niezbędnej do prawidłowego funkcjonowania procesu nauczania. Dzięki otwartej strukturze i modułowej budowie systemu, zrealizowana funkcjonalność stanowi doskonałą podstawę do dalszej rozbudowy programu, poszerzania go o nowe funkcje np. generator testów czy rozbudowany system nadzoru procesu nauczania.

Wykorzystanie i dalszy rozwój zrealizowanego systemu w znacznym stopniu zależą od aktualnych potrzeb i możliwości uczelni. O jego praktycznej przydatności będzie można powiedzieć wkrótce, po pewnym okresie jego wykorzystywania. System ten może być użyty jako wsparcie klasycznego procesu nauczania lub w realizacji studiów zdalnych.

BIBLIOGRAFIA

- [1] Kuszewski A., Łapienko K.: *Realizacja systemu zdalnego nauczania*, s. 25-41, Pro Dialog, nr 17, 2004.
- [2] Monsun-Haefel R.: *Enterprise Java Beans*, 3rd Edition, O'Reilly 2001.
- [3] Kuszewski A., Łapienko K.: *Zdalne nauczanie, realizacja systemu*, Instytut Informatyki PW, praca magisterska, luty 2004.

A CONCEPT OF AN e-LEARNING SYSTEM IN JAVA

Summary

A concept of an e-learning system implemented in the Institute of Computer Science Warsaw University of Technology is presented. System models are made in UML, in Rational Rose and the implementation is in Java, so it can execute in different environments. The client – server architecture was chosen. Remote Method Invocation was used to solve the communication problems. This is an asynchronical system in which the communication among users is through internal mail module. The functionality of this system comprises the functionality of electronic mail and file transfer protocol. The advantage of this system over email and ftp are strongly checked access permissions of each action executed in the system.

Anna Bobkowska, Maciej Piechówka, Stanisław Szejko

Katedra Inżynierii Oprogramowania, Politechnika Gdańska

**PROJEKT GRUPOWY W KSZTAŁCENIU STUDENTÓW
INFORMATYKI - DOŚWIADCZENIA NA WYDZIALE ETI
POLITECHNIKI GDAŃSKIEJ**

Streszczenie

Projekt grupowy stanowi uznaną, efektywną metodę kształcenia w zakresie inżynierii oprogramowania. W referacie omówiono założenia, organizację, kształt procesu projektowego, wzorce dokumentacyjne i wsparcie narzędziowe wypracowane w czasie kilkuletniej realizacji zajęć Projektu Grupowego na Wydziale ETI PG. Prezentowane wzorce i doświadczenia powinny okazać pomocne przy zamierzonym wprowadzeniu Projektu Grupowego do programu studiów.

1. WSTĘP

Projekt grupowy jest powszechnie uważany za efektywną metodę kształcenia w zakresie inżynierii oprogramowania. Stanowi on uwieńczenie edukacji z zakresu wytwarzania oprogramowania, dbałości o wymaganą jakość produktu, organizacji pracy, planowania i zarządzania przebiegiem projektu, a nade wszystko – narzuca integrację wiedzy zdobywanej w ramach odrębnych przedmiotów. Inne zadania studenckie zazwyczaj mają ograniczony zakres, dostosowany do konkretnego przedmiotu, np. na przedmiotach związanych programowaniem studenci poznają jedynie języki programowania, na przedmiotach związanych z analizą systemu – są nauczani głównie metod modelowania, itd. Koncentracja w ramach przedmiotów na wybranych zagadnieniach jest konieczna, jednak bez przedmiotu integrującego wiedzę studenci mają trudności z dostrzeżeniem powiązań pomiędzy treścią różnych przedmiotów. W ramach informatycznego projektu grupowego studenci są zmuszeni do praktycznego wykorzystania wiedzy z zakresu inżynierii wymagań, analizy i projektowania systemów, zarządzania projektem, jakości oprogramowania, programowania i testowania.

Bardzo istotne jest połączenie aspektów technicznych oraz organizacyjnych, a także fakt, że dzięki wpływowi na decyzje podczas całego cyklu wytwarzania oprogramowania, który rozciąga się od definicji problemu aż do wdrożenia (i obejmuje np. wybór podejścia, wybór strategii rozwiązania problemu, wybór architektury i konkretnych technik jakościowych), studenci rzeczywiście mogą zobaczyć konsekwencje swoich działań. Jest to szczególnie ważne w przypadku młodych ludzi bez większego doświadczenia zawodowego.

Projekt grupowy zbliża nas także do rzeczywistych projektów informatycznych, gdzie praca ma zazwyczaj charakter zespołowy, a nie indywidualny lub w parach (co na uczelniach wynika zwykle z konieczności oceniania pracy studenckiej). Istotną cechą jest praca „z rzeczywistym klientem”, który zazwyczaj ma swoje potrzeby, oczekiwania, ale także i ograniczenia. Projekt rozwija więc umiejętność współpracy i komunikacji interpersonalnej, uczy zarządzania zadaniami i ich współbieżnym wykonywaniem, zrozumienia występujących w projekcie ról oraz podziału zadań i odpowiedzialności.

Nauczanie pracy grupowej jest zalecane przez uznane organizacje standaryzujące i proponowane minima programowe kierunków informatycznych. M.in., w najnowszym programie ACM/IEEE wymienia się „pracę jako członka zespołu” wśród elementów nauczania profesjonalizmu w inżynierii oprogramowania [8], a wśród proponowanych przez Computing Curricula 2001 kursów są *CS292 Software Development and Professional Practice* lub zamiennie – jedno- bądź dwusemestralny projekt „wieńczący” nauczanie *CS 490 Capstone Project* [3]. Wiele uczelni wprowadza pracę grupową na informatycznych studiach II stopnia, na przykład jeden z najbardziej znanych uniwersytetów brytyjskich, Newcastle University, ma w programie dwusemestralny kurs *Software Engineering with Group Project* [6], o podwójnej liczbie punktów kredytowych.

Rangę pracy grupowej w kształceniu informatyków w Polsce potwierdza standard przyjęty przez Radę Główną Szkolnictwa Wyższego, który w ramach przedmiotów kierunkowych tak na studiach magisterskich jak i zawodowych przewiduje „Projekt. Indywidualny lub zespołowy, zaawansowany projekt programistyczny” [9].

Na Wydziale Elektroniki, Telekomunikacji i Informatyki Politechniki Gdańskiej Projekt Grupowy prowadzony jest od 1997 na specjalności Inżynieria Systemów i Bazy Danych. Idea jego wprowadzenia i pierwsze wdrożenie na WETI PG pojawiły się wraz z pracami programu Tempus SJEP-11137-96 *Industrial Software Development* [2,4] (stąd do dziś niektóre formularze dokumentów są w języku angielskim!). Kolejne lata przyniosły ulepszenie idei projektu i wykorzystywanych wzorców w oparciu o rozwiązania zaproponowane przez prof. Górskiego. W 1999 roku zespół wykładowców za wdrożenie projektu grupowego na WETI otrzymał Nagrodę I stopnia Rektora Politechniki Gdańskiej.

Zbliżony kształtem Projekt Grupowy wykonywany jest także na Studium Podyplomowym Zarządzania Projektami Telekomunikacyjnymi i Teleinformatycznymi, a w wersji ze „złotym rozwiązaniem” prowadzony jest na Studium Podyplomowym Inżynierii Oprogramowania.

Kolejne rozdziały prezentacji przedstawiają założenia przyjęte przy wprowadzaniu Projektu Grupowego na specjalności ISiBD, kształt przedmiotu i szablony dokumentów, wykorzystywane wsparcie narzędziowe oraz doświadczenia zdobyte w czasie kilkuletniej realizacji Projektu.

2. ZAŁOŻENIA ORGANIZACJI PROJEKTU

Projekt Grupowy został usytuowany na 8 semestrze specjalności ISiBD kierunku Informatyka, co pozwala wykorzystać wiedzę nabywaną przez studentów na wcześniejszych zajęciach kierunkowych, specjalnościowych i prowadzonych równolegle zajęciach z zakresu zarządzania projektami informatycznymi, a jednocześnie poprzedza semestry dyplomowe, nie kolidując z nimi obciążeniem. Poniżej przedstawiona jest lista głównych założeń organizacyjnych:

- Praca studentów w grupach 4 – 6 osób. Zasada jest samodzielne organizowanie się studentów w grupie roboczej jak i dowolny dobór infrastruktury komunikacyjnej (w zakresie oferowanym przez Wydział).
- Praca z zewnętrznym klientem i rzeczywistym problemem. Przykładami tematów mogą być *Teatr Edukacyjny* realizowany dla organizacji pozarządowej - Stowarzyszenia Teatralno-Edukacyjnego „Wybrzeżak”, *Edytor wiadomości MMS* dla Wirtualnej Polski SA., czy też *Prototyp narzędzia wspierającego metodę Characteristics – Questions – Metrics* jako projekt wspomagający badania naukowe realizowane na Wydziale. Studenci są stawiani przed problemem i ich zadaniem jest przewidywanie nakładu pracy oraz negocjowanie zakresu i kształtu projektu w celu dostosowania do dysponowanych zasobów. W praktyce poznają również problemy współpracy ze stawiającym wymagania i oceniającym produkt klientem. Dodatkowe aspekty to konieczność poznawania dziedziny problemu i komunikacji z odpowiednimi ekspertami oraz wymuszenie spojrzenia w szerszym kontekście, np. problemów prawnych lub etycznych. Akceptacja rozwiązania przez klienta i stopień jego usatysfakcjonowania wpływają na zaliczenie przedmiotu;
- Możliwie niewielka ingerencja prowadzącego zajęcia w pracę grupy, zapewniająca samodzielność i odczucie konsekwencji podejmowanych decyzji. Zakłada się, że to studenci wykazują inicjatywę w kontaktach z nauczycielem oraz że prowadzący, poza oceną końcowych rezultatów oraz przebiegu projektu, pełni rolę doradczą;
- Elastyczność doboru strategii wytwarzania. Wstępnie zakłada się pracę według cyklu klasycznego i do niego dostosowane są formularze i predefiniowane terminy realizacji poszczególnych etapów (zob. p. 3). Jednak, jeśli specyfika projektu lub prowadzone prace projektowe prowadzą do zmian strategii i/lub harmonogramu, studenci mają pełną możliwość zarządzania cyklem wytwórczym – przy spełnieniu wymogu uzyskania akceptacji nauczyciela prowadzącego projekt;
- Wytwarzanie w technologii obiektowej, choć i tu nie stanowi ono kanonu, od którego nie wolno odstępować;
- Pełna dokumentacja prac i produktów poszczególnych etapów pracy. Istotne jest zrozumienie roli dokumentacji (technicznej, zarządczej, planów) w projekcie i pozbycie się przekonania, że jest ona „niepotrzebnym sprawozdaniem”;
- Wsparcie dokumentacyjne i narzędziowe w zakresie wzorców dokumentów, strony przedmiotu, narzędzi CASE oraz systemu dostarczania i wersjonowania dokumentów.

3. KSZTAŁT PROCESU I SZABLONY PRODUKTÓW

Realizacja założeń o dużej dowolności w planowaniu (np. iteracyjny proces wytwarzania), jak również wysoka unikalność realizowanych projektów wprowadza wymagania na elastyczny opis podstawowych zasad realizacji projektów grupowych. Niezbędne informacje zawarte zostały w Przewodniku Studenta, który jest udostępniany studentom na początku zajęć. Zawiera on opis zasad przedmiotu i wymagań względem studentów, jak również szczegółowy opis zadań i produktów w poszczególnych fazach wytwarzania systemu. Fragment takiego opisu dla fazy planowania projektu i specyfikacji wymagań pokazany jest w tabeli 1.

Proces wytwarzania rozpoczyna się od czynności związanych z zarządzaniem projektem: planowania projektu i określenia infrastruktury (faza: planowanie projektu) oraz

określenia wizji i zakresu produktów (faza: studium wykonalności) w celu minimalizacji ryzyka niepowodzenia projektu.

Tabela 1

Fragment opisu procesu wytwarzania: etapy, zadania, produkty i terminy

<i>Etap</i>	<i>Tydz</i>	<i>Zadania</i>	<i>Produkty</i>	<i>Termin</i>
3. <i>Planowanie projektu</i>	3	<i>organizacja zespołu i podział ról planowanie projektu: faz, zadań, produktów poszczególnych faz, potrzebnych zasobów ustalenie harmonogramu projektu ryzyko i aspekty jakości wykonanie raportów</i>	<i>Plan projektu Plan fazy specyfikacji wymagań Raport fazy Raporty ze spotkań</i>	18.03. 2003
4. <i>Specyfikacja wymagań</i>	4-5	<i>Pozyskanie wymagań przy zastosowaniu metody odpowiadającej typowi projektu (analiza dokumentów, wywiady, ankietowanie, burze mózgów, prototypowanie) Zebranie pozyskanych wymagań Zapewnienie weryfikacji i akceptacji wymagań planowanie i akceptacja testów akceptacyjnych wykonanie czynności związanych z zapewnieniem jakości executing quality control activities wykonanie raportów</i>	<i>Specyfikacja wymagań systemowych Plan testów akceptacyjnych Plan fazy analizy Słownik danych Raport zmian Raport fazy Raporty ze spotkań</i>	01.04. 2003

Następne fazy mają charakter techniczny, natomiast czynności zarządzania i zapewnienia jakości są w nie wkomponowane. Czynności zarządzania dokumentowane są w każdej fazie za pomocą planu fazy, raportów ze wszystkich spotkań (z klientem i z nauczycielem) oraz raportu z fazy i raportu zmian (jeżeli takie miały miejsce). Dzięki dokładności opisu procesu projektowego możliwe jest stwierdzenie postępów w projekcie nawet w czasie, gdy nie ma spotkań z nauczycielem. Ponadto takie notatki ze spotkań są dla grup projektowych cennym zasobem wiedzy pomagającym śledzić źródło i rozwój konkretnych wymagań. Czynności zapewnienia jakości są dokumentowane jako raporty z inspekcji i wyniki testów. Fazy związane z wytwarzaniem oprogramowania obejmują:

- specyfikację wymagań, która stanowi punkt odniesienia przy odbiorze systemu;
- analizę systemu, w czasie której wykonywane są modele z zastosowaniem UML;
- projekt systemu, który zawiera dekompozycję systemu na podsystemy i inne decyzje strategiczne dla implementacji systemu;
- projekt podsystemów i obiektów;
- implementację w wybranym środowisku programistycznym;
- testowanie, w ramach którego przeprowadzane jest testowanie oraz usuwane są wykryte błędy.

Dla wszystkich dokumentów dostępne są na stronie przedmiotu szablony, z których korzystają studenci w czasie przygotowania dokumentów. Przykładowy szablon planu testów akceptacyjnych pokazany został na rysunku 1.

Nazwa projektu:

Wykonał:

Zespół projektowy:

Zatwierdził:

Faza: **Pozyskanie wymagań**

Nazwa dokumentu: **Plan testów akceptacyjnych**

Powiązane dokumenty:

Wersja:

Spis treści

Wprowadzenie

Plan testów

Elementy testów

Testowane właściwości

Kryteria przejścia testu

Kryteria zawieszenia i wznowienia

Środowisko

Projekt testu (dla każdego testu)

Identyfikator testu

Elementy testu

Testowane właściwości

Kryteria przejścia testu

Specyfikacja przypadków testowych (dla każdego testu)

Identyfikator przypadku testowego

Specyfikacja wejścia

Specyfikacja wyjścia

Specjalne wymagania proceduralne i środowiskowe

Zależności pomiędzy przypadkami testowymi

Procedura testowa (dla każdego testu)

Identyfikator procedury testowej

Specjalne

Kroki proceduralne

Raport testów (dla każdego wykonania procedury testowej)

Identyfikator raportu testu

Opis

Czynności i wejścia danych

Rys.1. Szablon planu testów akceptacyjnych

4. WSPARCIE NARZĘDZIOWE

W projekcie stosowane są następujące narzędzia wspomagające prace projektowe:

- Strona WWW Projektu Grupowego, która pełni funkcję informacyjną. Studenci mogą znaleźć na niej wszystkie informacje potrzebne do realizacji projektu oraz Przewodnik Studenta. Ze strony można także pobrać szablony dokumentacyjne;
- Profesjonalne narzędzia CASE i środowiska programistyczne;
- Lotus Notes, który jest narzędziem klasy CSCW i wspomaga dokumentowanie produktów według faz procesu wytwarzania.

Narzędzia komputerowego wspierania pracy grupowej (ang. *Computer Supported Cooperative Work* – CSCW) wspomagają współpracę pomiędzy fizycznie odległymi użytkownikami lub zespołami przy realizowaniu przedsięwzięć prowadzących do osiągnięcia wspólnego celu, realizowanych przy użyciu technologii informatycznych, która dostarcza infrastruktury do komunikacji, koordynacji i kooperacji oraz zarządzania dzielonymi zasobami. Wśród oferowanych produktów znaleźć możemy zarówno rozwiązania komercyjne jak i dostępne bezpłatnie [5]. Przedstawicielem pierwszej grupy jest produkt Lotus Domino, natomiast interesującym rozwiązaniem należącym do drugiej grupy narzędzi jest system BSCW.

W ramach prowadzonych projektów i prac dyplomowych opracowano kilka aplikacji wspomagających prowadzenie projektów grupowych [7], głównie opartych na środowisku Lotus Domino, które dostarcza większość usług wymaganych do budowy tego typu systemów. Wykonane aplikacje pozwalają na definiowanie i tworzenie różnych cykli wytwórczych oprogramowania, wspomagają zarządzanie repozytorium dokumentów i komunikację między uczestnikami projektu. Wykorzystywany w ramach projektu grupowego system składa się z dwóch podstawowych modułów: bazy danych dokumentów, oraz podsystemu zarządzania zadaniami wykonywanymi w ramach projektu. Baza danych umożliwia zarządzanie konfiguracją dokumentów, czyli pozwala na uporządkowane przechowywanie i chroniony dostęp do dowolnego dokumentu utworzonego w trakcie pracy nad systemem. Do każdego dokumentu dołączane są dodatkowe dane wpisywane automatycznie informujące o dacie i czasie umieszczenia dokumentu w bazie oraz pola informujące o fazie projektowej, do której dany dokument należy. Dodatkowo udostępniany jest mechanizm wersjonowania. System umożliwia też rejestrowanie przeglądów dokumentów przy wykorzystaniu różnych strategii. Przykładowy widok zawartości repozytorium projektowego przedstawia rys. 2.

Title	Creation	Author
▼ 1. Initial Project Planning		
List of teams and the assigned topics	2002-02-27	Maciej MP Piechówka
Initial Project Plan	2002-03-21	Wojciech Waloszek
▼ 2. Feasibility Study		
▼ 3. Project Planning		
Preliminary Project Plan	2002-04-25	Wojciech Waloszek
Preliminary Project Plan	2002-07-01	Wojciech Waloszek
Preliminary Project Plan	2002-07-01	Wojciech Waloszek
▼ 4. Elicitation of Requirements		
▼ 6. System Design		
System Design Documents (draft)	2002-06-17	Michał Zawadzki
Object Design and Implementation Plan	2002-06-17	Michał Zawadzki
Phase report	2002-07-01	Wojciech Waloszek
▼ 7. Object Design		
Object Design (draft)	2002-07-01	Wojciech Waloszek
Phase report	2002-07-01	Wojciech Waloszek
▼ 9. Finalising		
Documentation	2002-07-01	Wojciech Waloszek
Final Report (Project history, management and quality activities - discipline included)	2002-07-01	Wojciech Waloszek

Rys.2. Widok listy dokumentów zwartych w repozytorium projektu

Podsystem zarządzania zadaniami służy do integracji pracy zespołu w ramach projektu. Do tego celu wykorzystywane są mechanizmy oparte o pocztę elektroniczną, służące do komunikacji między członkami zespołu. Podsystem zarządzania zadaniami wspomaga czynności związane z planowaniem prac nad systemem w zakresie definiowania zadań, określania czasu ich trwania oraz czynności związane z kontrolą przestrzegania terminów.

Wykorzystywane narzędzie umożliwia zdalny dostęp do repozytorium tak przez studentów jak i wykładowców. Repozytorium stanowi także archiwum prowadzonych prac projektowych studentów.

5. DOŚWIADCZENIA

Projekt Grupowy w omawianym kształcie jest przez nas prowadzony od roku akademickiego 97/98 (z roczną przerwą) na studiach magisterskich i podyplomowych. Projekty obejmowały szeroki wachlarz oprogramowania: systemy dedykowane, prototypy narzędzi, systemy edukacyjne, systemy bazodanowe i aplikacje internetowe. Nie stwierdziliśmy problemów w wytwarzaniu aplikacji jakiegось konkretnego typu. Do najważniejszych doświadczeń i opinii na temat przebiegu projektu można zaliczyć:

- Z zasady prowadzone projekty miały charakter wytwórczy, występowały też zadania ponownej inżynierii działających aplikacji lub projekty stricte wdrożeniowe. O ile różne strategie wytwarzania są studentom znane, o tyle planowanie wdrożenia i utrzymania napotyka na spore trudności (widoczne też często w realizacji prac dyplomowych o podobnym charakterze).
- Bardzo dobre są rezultaty pracy studenckich grup projektowych, zarówno w sensie funkcjonalności i jakości dostarczanych produktów, jak i w zakresie współpracy z klientami zewnętrznymi, którzy zazwyczaj wyrażają satysfakcję ze współpracy ze studentami i zadowolenie z dostarczanych produktów.
- W kilku sytuacjach projekty przeciągnęły się ponad planowany na nie czas, ale tylko w jednym przypadku zadanie zostało zrealizowane dopiero w następnym roku.
- W kilku przypadkach rolę klienta pełnił prowadzący zajęcia (np. [1]), jednak obarczało to projekt łatwymi do przewidzenia minusami – brakiem rozróżnienia tych ról przez studentów (i prowadzącego!), nieformalnymi i *ad hoc* kontaktami z prowadzącym, rozmyciem fazy przekazania pracy.
- Obciążenie studentów Projektem jest bardzo duże. W aktualnym planie studiów jest to 75 godz. na studenta, natomiast praktyka pokazuje, że nie jest to ilość wystarczająca.
- Analogicznie, należy liczyć się z bardzo dużymi obciążeniami prowadzących zajęcia – szczególnie na początku i na końcu pracy nad projektem (i to, mimo, że pracujemy z bardzo dobrymi studentami). Należy szacować, że prowadzący jest w stanie dobrze prowadzić nie więcej niż dwie grupy projektowe w semestrze.
- Praca w grupach przebiega na ogół bez większych konfliktów, choć znane nam są sytuacje (nieplanowanych) zmian ról, w szczególności – kierownika projektu. W co najmniej jednym przypadku ostre wymagania klienta i napięte terminy spowodowały spore rozdziewięki w grupie na tle obciążenia pracą.
- Nie udało się też wprowadzić, podpatzonego na zachodnich uczelniach, wzajemnego oceniania pracy w obrębie grupy projektowej i podziału punktów przydzielanych *en bloc* przez nauczyciela i klienta. Nasi studenci kończyli zwykle te dyskusje dość burzliwie i z wnioskiem, by oceny wystawić "po równo" wszystkim członkom zespołu.

W anonimowych ankietach lat ubiegłych przeprowadzanych wśród studentów uczestniczących w projekcie przedmiot otrzymał wysokie opinie. Doceniono "lekcje" wyniesione z pracy z rzeczywistym klientem, doświadczenia pracy w grupie, konieczność posiadania całościowej wizji projektu. Narzekano głównie na obciążenie pracą i zbyt późne usytuowanie projektu w programie nauczania (to ostatnie podnosili przede wszystkim studenci pracujący, którzy wcześniej odczuwali brak praktyki projektowej). W Katedrze Inżynierii Oprogramowania oceniamy, że wprowadzenie Projektu Grupowego istotnie podniosło jakość realizowanych na specjalności prac dyplomowych.

6. ZAKOŃCZENIE

Komisja Programowa kierunku Informatyka działająca przy Dziekanie Wydziału ETI doceniła rezultaty prowadzonych zajęć Projektu Grupowego i zaproponowała, by zajęcia te, w wymiarze 120 godzin, rozciągnąć na wszystkich studentów kierunku Informatyka. Wobec pozytywnej decyzji Rady Wydziału ETI i wprowadzenia zajęć Projektu Grupowego do programu wszystkich specjalności, autorzy mają nadzieję, że prezentowane tu doświadczenia okażą się przydatne w ich przygotowaniu i wdrożeniu.

BIBLIOGRAFIA

- [1] Barylski M., Olszewski M., Pardo K., Szczerkowski M.: *Prototyp narzędzia wspierającego metodę Characteristics – Questions – Metrics*. Dokumentacja zajęć Projektu Grupowego 2001/2002, KZI WETI PG
- [2] Bobkowska A.: *Konstruowanie Oprogramowania Równoległego*. Materiały wykładowe Projektu Tempus SJEP-11137-96. WETI PG, 1997
- [3] Computing Curricula 2001, ch. 9, Completing the Curriculum, ACM/IEEE Comp. Society.
<http://www.computer.org/education/cc2001/final/chapter09.htm>
- [4] Czaja J., Gromadzki T., Piechówka M. i Szejko S.: *Group Project Module*, Materiały wykładowe Projektu Tempus SJEP-11137-96. WETI PG, 1997
- [5] Lotus, IBM; SharePoint, Microsoft; BSCW - <http://bscw.fit.fraunhofer.de/>, CSCW
<http://www.usabilityfirst.com/groupware/cscw.tlx>
- [6] MSc in Computing Science Curriculum. University of Newcastle, School of Computing Science.
http://www.cs.ncl.ac.uk/degrees/pg/compsci/2003compsci.htm#_Toc46896936
- [7] Sawczuk L., Szwaczyk D., *System wspomagania projektu grupowego*, 1999; Tatara Sz.: *Zarządzanie projektem informatycznym w oparciu o WWW*, 1999; Landowski P., *System wspomagający zarządzanie projektem informatycznym*, 2001; Adamkiewicz D., *System wspierania pracy grupowej w Studium Podyplomowym Inżynierii Oprogramowania*, 2002.
- [8] Sobel A.E.K (ed): *Final Draft of the Software Engineering Education Knowledge*. Computing Curricula – Software Engineering Volume. ACM/IEEE Comp. Society, <http://sites.computer.org/ccse/>
- [9] *Standardy nauczania dla kierunku studiów Informatyka*. Załącznik 12, MENiS, 2002.

GROUP PROJECT IN EDUCATION OF SOFTWARE ENGINEERING – TEACHING EXPERIENCE AT GDANSK UNIVERSITY OF TECHNOLOGY

Summary

Group project is an effective method in software engineering education. This paper presents the objectives of Group Project course, supervision aspects, the structure of development process, product templates, supporting tools and experience gained while teaching Group Project classes at Faculty of Electronics, Telecommunication and Informatics of the Gdansk University of Technology.

Bogdan A. Galwas, Remigiusz J. Rak

Ośrodek Kształcenia na Odległość, Politechnika Warszawska

EDUKACJA W SPOŁECZEŃSTWIE INFORMACYJNYM

Streszczenie

Rozwój technologii informacyjnej stawia przed nami trudne wyzwania, ale też technologia ta wyposaża nas w nowe i nieznane do tej pory narzędzia, które już wykazują wielką przydatność w procesie kształcenia. Komputer i Internet stają się niezastąpionymi narzędziami pracy wykładowcy i studenta. Płyty CD stają się nośnikiem nowej generacji podręczników multimedialnych w formacie HTML, w których tradycyjny tekst, ilustracje i rysunki uzupełnione zostały specjalnymi technikami animacji i prezentacji, wmontowanymi komentarzami audio i video, symulacjami obliczeń i eksperymentów. Rodzi się nowa jakość: zdalny dostęp do laboratorium oraz laboratorium wirtualne.

1. WSTĘP

Żyjemy w czasach wielkich i szybkich zmian, zachodzących w wymiarze krajów, kontynentów i globu. Dostrzegamy z łatwością, że siłą napędową tych zmian jest lawinowy rozwój technologii. Szczególnie ważną rolę ogrywa rozwój technologii informatycznej, telekomunikacji, technik multimedialnych i Internetu. Społeczeństwo, w którym wymienione czynniki odgrywają istotną rolę nazywamy informacyjnym.

Wielu z nas z entuzjazmem wita skutki tego rozwoju, ale wielu patrzy z troską i obawą na rosnące bezrobocie, skażenie środowiska, możliwość utraty zawodu i niepewność jutra. Lekarstwem, prawdopodobnie najważniejszym, jest edukacja, która pozwoli lepiej zrozumieć otaczający świat i przystosować do zmieniających się warunków pracy i życia. Społeczeństwa zrozumiały, że wykształcenie musi być dostępne dla wszystkich, że proces kształcenia nie kończy się z uzyskaniem dyplomu szkoły wyższej, a musi być kontynuowany przez cały okres aktywności zawodowej i biologicznej. Dalej, że trzeba ponieść koszty systemu, który to zapewni. Przystępujemy więc, do budowy systemu "Kształcenia przez całe życie", w którym dobrze rozwinięte i wyposażone szkolnictwo podstawowe, średnie i wyższe zostanie uzupełnione siecią ośrodków kształcenia ustawicznego i uniwersytetów otwartych. Istnieje możliwość i potrzeba opracowania nowoczesnych narzędzi edukacji. Należy do nich między innymi elektroniczny podręcznik multimedialny. Rodzi się potrzeba opracowania metod symulacji eksperymentu, wirtualnego laboratorium oraz zdalnego dostępu do laboratorium rzeczywistego.

2. STRUKTURA PODRĘCZNIKA MULTIMEDIALNEGO

Tradycyjny podręcznik jest książką wydrukowaną na papierze, napisaną i zredagowaną w sposób ułatwiający studentowi zrozumienie materiału i osiągnięcie umiejętności posługiwania się opanowaną wiedzą. Do jego przygotowania wykorzystuje się technologię doskonałą przez ostatnie pięćset lat. Ale oto w latach 90-tych, pojawia się nowa technologia zapisywania tekstu, rysunków i wzorów, nowy sposób przechowywania tak przygotowanych materiałów na dyskietkach, twardych dyskach, płytach CD i DVD, na serwerach portali edukacyjnych, w bibliotekach elektronicznych. Materiał ten można drukować, ale można też wyświetlić go na ekranie monitora, czy też ekranie sali wykładowej. Tak powstały książki elektroniczne, zwane też książkami multimedialnymi [2,7].

Nowa forma podręczników, nowa nie tylko ze względu na sposób zapisywania materiału, jest właśnie intensywnie rozwijana i doskonalona. Rozwój dokonuje się w dwu kierunkach. Jeden kierunek to absorpcja nowych technik prezentacji, animacji i symulacji. Drugi kierunek związany jest z rozwojem modelu zapisywania, organizacji i przekazywania wiedzy [1,3,6].

W latach 2001/02 w Ośrodku Kształcenia na Odległość Politechniki Warszawskiej (OKNO PW) opracowano nowy, gałęziowy model podręcznika multimedialnego [4,5]. Zgodnie z założeniem podręcznik taki ma zastąpić nie tylko tradycyjną książkę drukowaną, ale wykład. Nie podlega dyskusji, że ten nowy model podręcznika jest szczególnie użyteczny dla studentów kształcących się w trybie „na odległość” [4,5,7].

2.1. Charakterystyka podręcznika akademickiego

Elektroniczny podręcznik akademicki winien spełniać szereg wymagań. Uniwersalnym wymogiem, jak w przypadku każdego podręcznika jest to, aby był napisany i zredagowany w sposób ułatwiający studentowi zrozumienie materiału i osiągnięcie umiejętności posługiwania się opanowaną wiedzą. Warunek ten jest zwykle spełniony przez autorów w mniejszym, lub większym stopniu, gdyż napisanie dobrego podręcznika jest sztuką samą w sobie.

Kolejnym wymóg to użycie odpowiednich narzędzi. Tradycyjnie drukowane podręczniki książkowe wykorzystują dobrze znane narzędzia technologii druku. Są to: tekst z wykorzystaniem rozmaitych fontów, pismo pogrubione, pochyle, czy też kolorowe, wzory matematyczne i równania, rysunki rozmaitych typów, zdjęcia, dobór koloru tła itp. Wszystkie wymienione narzędzia powinny być rozsądnie i celowo użyte.

Oddzielnym problemem jest celowe użycie całego szeregu nowych narzędzi oferowanych przez techniki multimedialne. Można tu wymienić: komentarze pisane, komentarze mówione (audio), komentarze jako pliki filmowe z głosem (video & audio), animacje rysunków, animacja prezentacji PowerPoint z komentarzem wykładowcy, generator testów, wreszcie symulacje obliczeń ilustrujących wywody teoretyczne, czy wręcz symulacje eksperymentu.

Bardzo ważną rolę odgrywa tutaj język Java. *Aplety Javy* mogą być wplecione do podręcznika przygotowanego w formacie HTML, mogą go „ożywić” i udoskonalić. Java potrafi wzbogacić multimedialną zawartość stron WWW oferując płynne animacje, zaawansowaną grafikę, dźwięk i video bez potrzeby zaopatrywania się w dodatkowe aplikacje i podłączanie ich do przeglądarki WWW. Skompilowany program może zostać

dystrybuowany w polimorficznej sieci Internet, bez potrzeby troszczenia się o to, jakiego systemu operacyjnego i jakiej przeglądarki używają odbiorcy.

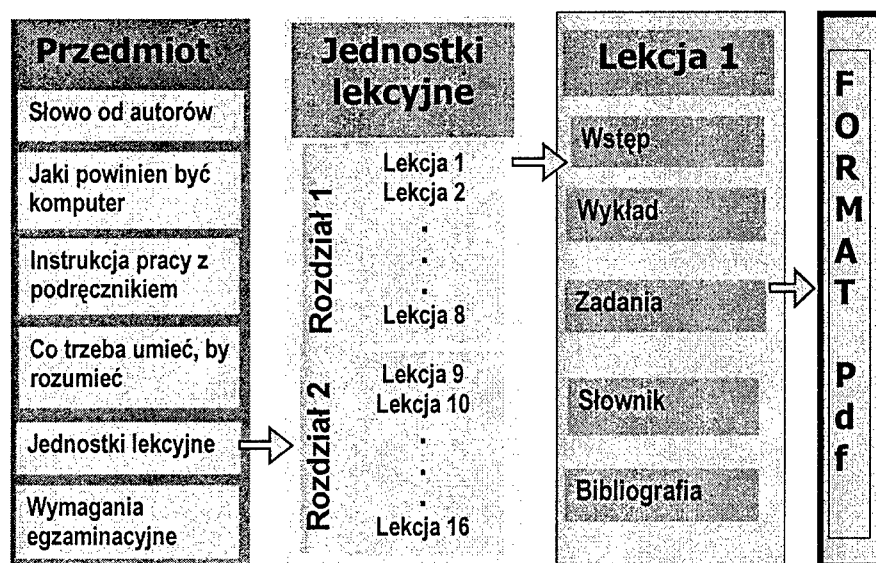
Lista dostępnych narzędzi multimedialnych jest długa, a same narzędzia są nieznane autorom tradycyjnych książek (wyjątkiem mogą być testy).

2.2. Gałęziowy model podręcznika elektronicznego

Struktura książki drukowanej nazywana jest **liniową** [1,5], choć bardziej właściwym wydaje się określenie **struktura szeregową**. Charakteryzuje się ona tym, że kolejne elementy następują po sobie, tekst, rysunki, równania, przykłady, bibliografia, dyskusja rozwiązań, w określonym przez autora porządku i proporcjach. Wyjątkiem są przypisy, umieszczone na dole strony, oraz wyodrębniona i umieszczona na końcu grupa „Dodatki”, czy skorowidz. Forma książki wymusza jej szeregową strukturę i jednowymiarowość narracji.

Podręcznik multimedialny, wykorzystujący tzw. **hypertext** stwarza możliwość napisania wykładu o **strukturze gałęziowej**. W opracowanym przez OKNO PW modelu podręcznika narracja **wykładu głównego** prowadzona jest jako obszerne streszczenie, zaopatrzone w wybrane, ale nieliczne przykłady. Specjalnie przygotowane „linki” prowadzą studenta do dodatkowych komentarzy: pisanych i dźwiękowych, do szczegółowych wyprowadzeń, do szczegółowych prezentacji w PowerPoint z obszerną argumentacją audio, do generatora testów, do przykładów zastosowań, analogicznych rozwiązań, itd..

Uproszczona struktura podręcznika multimedialnego przedstawiona jest na rysunku 1.



Rys.1. Uproszczona struktura podręcznika multimedialnego

Wykład główny posługuje się zwykle dwoma stylami. Powszechnie używany jest styl narracji [3] (opowiedz im o tym ...), ale wykładowcy chętnie też używają przykładów (pokaż im to ...).

Tradycyjne narzędzia **wykładu głównego** są zasadniczo identyczne z narzędziami książki drukowanej i zostały wymienione wcześniej. Autorzy posługują się zwykle: tekstem z różnymi wymiarami fontów, wzorami, rysunkami, zdjęciami, dobierając odpowiednio ich kolory i kolor tła. Autorzy tekstów drukowanych są ograniczani przez redakcje w obfitym używaniu wymienionych narzędzi, ze względu na koszty. Tych ograniczeń nie mają podręczniki elektroniczne. Ograniczenia wywołują jedynie zdrowy rozsądek i obawa przed pstrokazizną.

Nowe narzędzia otwierają nowe możliwości redagowania podręczników o nieznanych do tej pory cechach. Podstawowa wiedza **wykładu głównego (Jednostki Lekcyjne)** podana jest w kolejnych **segmentach** i ewentualnie w **punktach**. Student, studiując materiał dydaktyczny **wykładu głównego** napotyka szereg gałęziowo dołączonych plików, których otwarcie prowadzi go do dodatkowych komentarzy, ilustracji rysunkowych i animacji, do szczegółowych wyprowadzeń, do małych programów symulacyjnych, itp. Każdy z nich ułatwia mu zapoznanie się z głównymi тезami wykładu, pomaga w ich zrozumieniu.

W ramach jednego segmentu wykorzystana zwykle zostaje jedynie część narzędzi multimedialnych, zgodnie z logiką wykładu ustaloną przez autora. Nie ulega wątpliwości, że właściwego wykorzystania narzędzi multimedialnych trzeba się nauczyć.

Struktura gałęziowa wykładu ma wiele sprawdzonych zalet. Autorom ułatwia wykorzystanie rozmaitych narzędzi multimedialnych i funkcjonalnych w uzyskaniu dobrej prezentacji merytorycznych treści przedmiotu. Studentom ułatwia zrozumienie wykładu, zostawia wiele możliwości poruszania się po materiale, ułatwia powtórzenia, poszukiwanie konkretnych informacji, rozbudza ich aktywność. Prowadzone badania nad tempem przyswajania materiału z podręcznika szeregowego (liniowego) i gałęziowego, w oczywisty sposób, wykazały przewagę tego drugiego.

3. WIRTUALNE LABORATORIUM

Wirtualne przyrządy pomiarowe, w dużym stopniu zastąpiły w laboratoriach dydaktycznych drogie i skomplikowane przyrządy inteligentne, co znacznie uprościło proces projektowania, uruchamiania i modernizacji tych laboratoriów. Graficzny interfejs użytkownika (GUI), który do złudzenia przypomina rzeczywisty przyrząd pomiarowy powoduje, że użycie i rozumienie przyrządu jest intuicyjne dla tych, którzy korzystali do tej pory z przyrządów konwencjonalnych. Możliwość modyfikowania procedury pomiarowej poprzez zmianę zainstalowanego w komputerze oprogramowania, bez zmiany komponentów sprzętowych sprawia, że badania i eksperymenty stają się coraz bardziej elastyczne, nowoczesne i proste.

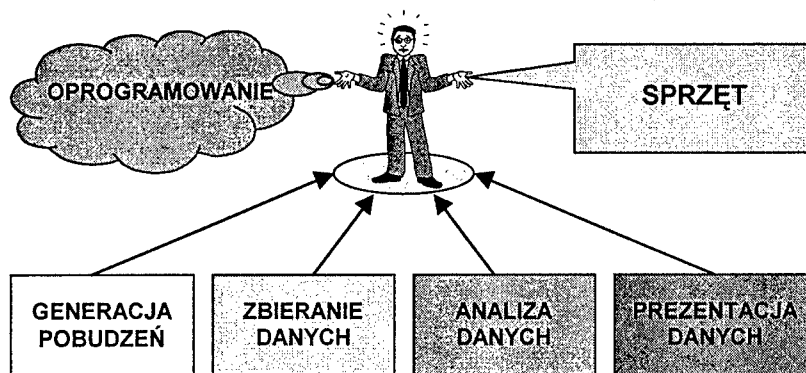
3.1. Idea przyrządu wirtualnego

Przyrząd wirtualny to rodzaj inteligentnego przyrządu pomiarowego powstałego w wyniku sprzężenia pewnego sprzętu nowej generacji z komputerem osobistym ogólnego przeznaczenia i przyjaznym dla użytkownika oprogramowaniem, które umożliwia użytkownikowi współpracę z komputerem na zasadach takich jakby obsługiwał tradycyjny przyrząd pomiarowy. Ideę przyrządu wirtualnego zilustrowano na rys.2. Blok decyzyjny oznaczono symbolicznie znakiem postaci projektanta, którym w pewnych przypadkach może stać się sam przyszły użytkownik [8,9].

Już sama definicja przyrządu wirtualnego, z uwagi na brak precyzji w sformułowaniach, wskazuje na dużą różnorodność w zakresie architektury przyrządów wirtualnych.

Z grubsza można je podzielić na trzy kategorie:

- Kat. A. Fizycznie istniejące przyrządy autonomiczne wyposażone w interfejsy przyrządowe IEC-625 lub RS232 (i pochodne), panel graficzny na ekranie monitora (symulujący płytę czołową) - obsługa przyrządu za pomocą „myszy”,
- Kat. B. Karta DAQ lub moduły VXI, (PXI, PCI - bez płyty czołowej, w miejsce przyrządu autonomicznego), panel graficzny na ekranie monitora (symulujący płytę czołową) - obsługa przyrządu za pomocą „myszy”,
- Kat. C. Brak fizycznego przyrządu (sprzętu – poza PC), dane wejściowe pobierane z plików w pamięci masowej, bazy danych, innych komputerów lub generowane w sposób numeryczny, panel graficzny na ekranie monitora (symulujący płytę czołową), obsługa za pomocą „myszy”.



Rys.2. Idea wirtualnego przyrządu pomiarowego

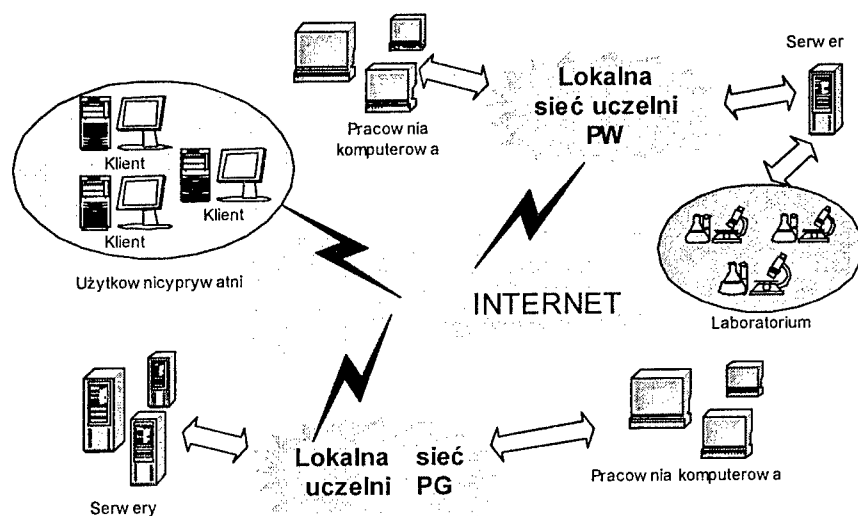
Łatwo zauważyć, że ostatnia kategoria dotyczy raczej symulacji przyrządu, systemu lub procesu pomiarowego i jest niezwykle przydatna w dydaktyce. Charakteryzuje się dużą uniwersalnością, elastycznością oraz niskim kosztem opracowania. Wyjątkowo dobrze nadaje się jako uzupełnienie do nowoczesnych podręczników szkolnych i akademickich wydawanych w formie elektronicznej. Wszystkie standardowe pakiety oprogramowania narzędziowego jak LabWindows/CVI, LabVIEW (*National Instruments*) i wiele innych umożliwiają projektowanie takich przyrządów w formie aplikacji pod Windows. Najbardziej jednak korzystne aplikacje to wspomniane już aplety Javy.

3.2. Charakterystyka wirtualnego laboratorium

Idea *wirtualnego przyrządu pomiarowego* może być w prosty sposób rozszerzona na laboratorium. Pojęcie *laboratorium wirtualne* staje się coraz bardziej powszechne i akceptowane, a w tradycyjnym rozumieniu, wiąże się z określeniem wirtualnej rzeczywistości, czyli kreowanego przez komputer fragmentu otaczającego nas świata.

Z drugiej strony niezwykle proste jest sprzęgnięcie, zdefiniowanego wcześniej, przyrządu wirtualnego (zawierającego sprzęt i oprogramowanie) z lokalną siecią komputerową. Wymaga tylko zainstalowania karty sieciowej w komputerze (kontrolerze) i przydzielenia mu numeru IP. Obsługa protokołu sieciowego, natomiast wbudowana jest do większości bibliotek oprogramowania narzędziowego. Rodzi to możliwość realizacji publikacji w sieci wyników eksperymentu, a w dalszej konsekwencji *zdalnego dostępu do laboratorium rzeczywistego*. Idea powstała na bazie laboratorium rozproszonego i coraz częściej dla ujednolicenia podejścia w sensie dydaktycznym, również określana jest mianem *laboratorium wirtualnego*.

Osadzenie wirtualnego przyrządu pomiarowego w rozproszonym systemie zlokalizowanym w sieci Internet daje niespotykaną do tej pory możliwość tworzenia zaawansowanych i elastycznych systemów, które mogą służyć prowadzeniu eksperymentów zdalnych i wspomagać proces dydaktyki. Szybki rozwój narzędzi programistycznych ułatwiających komunikację komputerów na duże odległości przesądza o wyjątkowej atrakcyjności wirtualnych laboratoriów. Możliwe staje się prowadzenie eksperymentów, oferowanych przez różne ośrodki naukowe, oraz korzystanie z wyników przez szerokie grono naukowców i studentów, niezależnie od miejsca ich aktualnego pobytu. Sprawia to, że wirtualne laboratoria pomiarowe są w ostatnich latach przedmiotem badań wielu instytucji naukowych. Idea laboratorium wirtualnego osadzonego w sieci Internet przedstawiona jest na rys. 3.



Rys.3. Schemat poglądowy wirtualnego laboratorium pomiarowego

Ośrodek badawczy, który udostępnia laboratorium (niech to będzie uczelnia PW), w sieci lokalnej stawia do dyspozycji swoich pracowników naukowych i studentów wszystkie bądź też wybrane, zasoby (przyrządy wirtualne). Po sprzężeniu sieci lokalnej z Internetem z zasobów laboratorium korzystać może inna uczelnia, np. PG. Korzyści z takiego stanu rzeczy czerpią obydwie ośrodki, ponieważ uczelnia PG może udostępniać uczelni PW wolną moc obliczeniową. W przypadku, gdyby uczelnia PG, na zasadzie

specjalizacji w innej dziedzinie, udostępniła swoje laboratorium wirtualne, obydwie uczelnie byłyby w stanie uzupełniać się w jeszcze szerszej skali. Możliwość korzystania z zasobów innych ośrodków badawczych jest szczególnie pożądana tam, gdzie chodzi o kosztowny, wysoko specjalistyczny sprzęt, który często jest poza zasięgiem wielu mniejszych uczelni i instytutów. W zasadzie, za pośrednictwem Internetu, z zasobów wirtualnego laboratorium mogą korzystać też inni upoważnieni do tego użytkownicy indywidualni, znajdujący się w swoich domach, szkołach, miejscach pracy itp. Upowszechnienie tego typu technologii ułatwi dostęp do wiedzy i umożliwi bardziej efektywne wykorzystanie potencjału ośrodków naukowo-badawczych. Przełamanie barier w prowadzeniu eksperymentów wykonywanych w trybie „on-line”, z możliwością bezpośredniego wpływania na badany proces poprzez dobór warunków pomiaru, źródeł i parametrów wymuszeń sprawią, że student będzie nie tylko biernym obserwatorem, ale świadomym uczestnikiem prowadzonych doświadczeń.

Dostęp do wirtualnego laboratorium nie musi być ograniczony do kilku godzin w tygodniu, lecz możliwy nawet przez okrągłą dobę. Jak już wcześniej wspomniano, jest to nieocenione zwłaszcza w przypadku drogiego sprzętu, do którego studenci mają ograniczony dostęp w laboratorium tradycyjnym. Studenci z każdego poziomu studiów mogliby korzystać zarówno z kosztownego sprzętu, jak i drogich obiektów badań, czy wręcz rozwijać indywidualne zainteresowania. Organizacja tego dostępu, dokonywanie autoryzacji oraz identyfikacji użytkowników stanowi odrębny problem, który jest już rozwiązany na pewnym poziomie w innych grupach zastosowań informatycznych.

5. ZAKOŃCZENIE

Książka elektroniczna jest tworem nowym, produktem rozwoju technik informatycznych i technologii informacyjnej. W pierwszym okresie traktowano ją jako inny sposób zapisu i przechowywania stron książek drukowanych. Wraz ze zdobywaniem doświadczeń w posługiwaniu się tym nowym narzędziem odkryto, że książkę elektroniczną, zawierającą materiały dydaktyczne dla studentów, można wyposażać w dodatkowe zbiory informacji drukowanych, mówionych i obrazowych, w ten sposób, aby ułatwić studentom przyswojenie i zrozumienie materiału. Nowy model podręcznika powstał z myślą o studentach studiów przez Internet, aby zastąpić im nie tylko podręcznik drukowany, ale i wykład.

Próba zastąpienia wykładu specjalnie przygotowanym podręcznikiem może nasuwać prosty wniosek, że najlepszym rozwiązaniem jest zarejestrowanie na taśmach video prowadzonych „na żywo” wykładów akademickich. Wartość takich rejestracji jest rzeczywiście duża, jednakże nie tak duża, jak można by w pierwszej chwili oczekiwać. Znowu powraca konieczność uzupełnienia wykładów książkami i stworzenie mechanizmów interaktywności, a to czyni problem prowadzenia przedmiotu dalej otwartym.

Próbując przewidzieć jak w najbliższym czasie rozwinię się technologia książek elektronicznych można wskazać na dwa elementy. W krótkim okresie kilku lat wykładowcy otrzymają nowe narzędzia, które umożliwią wprowadzenie do *Jednostek Lekcyjnych* eksperymentów symulowanych i prowadzonych na odległość eksperymentów rzeczywistych. Potaniecie i upowszechnienie technologii przygotowania dysków DVD umożliwi znaczne zwiększenie rozmiarów materiałów dydaktycznych, co w praktyce sprowadzi się do zwiększenia rozmiarów wmontowanych do wykładu plików video.

Integracja sieci komputerowych z laboratoriami rzeczywistymi daje możliwość tworzenia wielopoziomowych, i rozbudowanych struktur informacyjnych. Istniejące opro-

gramowanie internetowe, może być w łatwy sposób wykorzystane do organizacji migracji danych pomiędzy różnymi obiektami sieci. Standardowe języki programowania jak C++ i Java mogą być wykorzystane w narzędziowych środowiskach programistycznych do projektowania aplikacji sieciowych i internetowych. Protokoły internetowe, technologia ethernetowa oraz gniazdka (*sockets*) to elementy, które na tym etapie powszechnie wykorzystywane będą do projektowania struktury sieciowej.

Przekształcając społeczeństwo przemysłowe w informacyjne musimy oprzeć się na edukacji, nie tylko dlatego, że zmieni nam ona punkt widzenia na otaczający świat i pozwoli dostrzec grożące niebezpieczeństwa, ale także dlatego, że pozwoli nam zmieniać ten świat bez wielu pułapek i zagrożeń.

BIBLIOGRAFIA

- [1] Chou C., *Developing Hypertext-Based Learning Courseware for Computer Networks: The Macro and Micro Stages*, IEEE Trans. Edu., vol.42, pp. 39-44, February 1999.
- [2] Kaskinen T., Lacruz Chiva J. V., Climent M. M., Monforte Castello C., Montesinos Sanchis P., *The Great Paella Cookbook for Online Learning*, Universidad Politecnica de Valencia, CERES, 1999.
- [3] Janicki T., Liegle J.O., *Development and Evaluation of a Framework for Creating Web-Based Learning Modules: A Pedagogical and System Perspective*, Journal of Asynchronous Learning Network, Vol.5, Issue 1, May 2001, pp. 58-84.
- [4] Galwas B. A., Nowak J., Pajer M., Witonski P.: *New Model of Electronic-Book for Distance-Learning Courses*, Proceedings of 2001 EDEN Annual Conference, 15-18 June, Stockholm, 2001.
- [5] Galwas B., Barczyk J., Nowak S., Piwowarska E., Rak R.: *Web-based model of Engineering Studies Developed by Warsaw University of Technology*, Proceedings of European Conference „The New Educational Benefits of ICT”, 2-4 September, Rotterdam, 2002.
- [6] Khalifa M., Lam R., *Web-Based Learning: Effects on Learning Process and Outcome*, IEEE Trans. Edu., vol.45, pp. 350-356, November 2002.
- [7] Bielawski L., Metcalf D., *Blended eLearning*, HRD Press Inc., Amherst, Massachusetts, Chapter 12, 2003.
- [8] Rak R.J., Galwas B., *Virtual Laboratory - A Future Part of The New Web-Based Model of Undergraduate Engineering Studies Developed By Warsaw University of Technology*, Joint IMEKO TC-1 & XXXIV MKM Conference 2002, 8 - 12 September 2002, Wrocław.
- [9] Rak R.J., Pyszlak P., *The design of a Web-Based Virtual Laboratory – selected problems*, World Congress IMEKO, June 2003, Dubrovnik, Croatia.

EDUCATION IN THE INFORMATION SOCIETY

Summary

ICT - Information and communication technology is the source of the new tasks for the people, but at the same time it equips them in the new, sometimes unknown, tools which can be very helpful in the area of education process. Computer and the Internet are the irreplaceable tools for both, the lecturer and the student. The CD ROM's are the carriers of the new generation electronic books prepared in HTML format. The e-books besides the traditional text and figures are equipped with special techniques of presentation and animation, audio and video comments and experiment simulation. There is coming quite a new quality: remote access to laboratory and virtual laboratory.

Krzysztof Gierłowski, Krzysztof Nowicki

Katedra Systemów Informacyjnych, Politechnika Gdańska

ZASTOSOWANIE TECHNOLOGII INTERNETOWYCH W REALIZACJI SYSTEMÓW ZDALNEGO NAUCZANIA

Streszczenie

W referacie dokonano przeglądu systemów zdalnego nauczania. Przedstawiono ich podstawową klasyfikację oraz prowadzone prace standaryzacyjne, w szczególności standard SCORM. Omówiono technologie internetowe stosowane w systemach e-learningu oraz ich wpływ na sposób tworzenia tych systemów oraz efektywność procesu dydaktycznego. Wskazano najbardziej prawdopodobne kierunki rozwoju systemów e-learningu wspartych technologiami internetowymi.

1. WSTĘP

W okresie ostatnich kilku lat możemy obserwować znaczący rozwój idei zdalnej edukacji (e-learning), a w szczególności edukacji wspartej wykorzystaniem mediów elektro-nicznych. Rosnąca popularność różnych form zdalnego nauczania wynika z faktu, iż pozwalają one na znaczące zmniejszenie wpływu trzech podstawowych barier, utrudniających znacznej części potencjalnych odbiorców skorzystanie z tradycyjnej oferty edukacyjnej. Bariery te to czas, miejsce i koszt szkolenia.

Systemy zdalnego nauczania posługujące się technologiami takimi jak radio, telewizja, analogowy zapis obrazu oraz dźwięku i podobnymi, funkcjonują już od wielu lat i dobrze sprawdzają się w określonych środowiskach pracy. Jednakże dopiero zastosowanie sprzętu oraz sieci komputerowych (w szczególności ogólnosięciowego Internetu) w celach edukacyjnych pozwoliło tej metodzie nauczania stać się rozwiązaniem konkurencyjnym dla nauczania klasycznego. Nie jest to bynajmniej kres możliwości e-learningu, gdyż dzięki stopniowemu wdrażaniu systemów ITS (*Intelligent Tutoring System*) ma on szansę przewyższyć efektywnością jednoosobowe szkolenia indywidualne, przy zachowaniu swej olbrzymiej przewagi ekonomicznej [1,2].

W chwili obecnej olbrzymi wpływ na sposoby i systemy zdalnego nauczania mają technologie internetowe, oczywiście rozumiane nie tylko jako sposób przesyłania danych za pomocą sieci komputerowych, ale również jako cała gama sposobów:

- wprowadzania i zapisu,
- przechowywania,

- łatwego i elastycznego katalogowania i wyszukiwania,
- odczytu i przejrzystej prezentacji,

informacji wszelkich możliwych typów (tekst, obraz, dźwięk itp.) i to w sposób jak najbardziej naturalny i mało kłopotliwy dla przeciętnego użytkownika.

Technologie te odnoszą się wprost do podstawowych problemów, z którymi mamy do czynienia w procesie dydaktycznym. Spoglądając na sprawę z innej perspektywy, możemy uznać sieć Internet za narzędzie służące zdobywaniu informacji, a zatem za narzędzie edukacyjne.

2. PRZEGLĄD SYSTEMÓW ZDALNEGO NAUCZANIA

Wraz z rozwojem oraz powiększaniem się oferty produktów służących zdalnej edukacji, obserwujemy wyodrębnianie się kilku wyraźnych grup tych produktów, przeznaczonych do różnych zadań i dla różnych typów użytkowników.

2.1. Produkty służące budowie prezentacji (*Authoring tools*)

Przytłaczająca większość produktów przeznaczonych do obsługi zdalnego nauczania okazuje się, po prostu, narzędziami do tworzenia prezentacji i publikacji elektronicznych. Wbrew pozorowi prostocie zadania, realizowanego, ze zmiennym powodzeniem, powodzeniem od wielu lat, ma ono kluczowe znaczenie dla popularyzacji e-learningu, gdyż decyduje o wzroście bazy tematycznej, którą obejmują elektroniczne szkolenia. Brak łatwych w obsłudze narzędzi tego rodzaju przyczynił się w przeszłości do zahamowania rozwoju nauczania elektronicznego na poziomie wstępnym i w dziedzinach nie związanych z informatyką – powodem były trudności w tworzeniu kursów przez personel dydaktyczny bez dodatkowego przygotowania informatycznego.

Narzędzia budowy prezentacji stosują znane od lat technologie zapisu i prezentacji statycznych dokumentów wykorzystywane w sieci Internet (jak języki HTML, FLASH, Schockwave, Postscript, oraz metody kompresji i zapisu danych multimedialnych), a opisane pokrótce w dalszych rozdziałach.

W chwili obecnej jednak, proste publikowanie statycznego materiału dydaktycznego (nawet rozszerzonego o materiał audio i video) w formie elektronicznej w małym jedynie stopniu wykorzystuje oferowane nam przez nowoczesne technologie możliwości. Dlatego bardziej zaawansowane produkty tej grupy, oferują znacząco większe możliwości, dzięki zastosowaniu języków skryptowych, języków programowania niezależnych od platformy sprzętowo/programowej oraz klasycznych programów dołączanych do prezentacji.

Technologie te pozwalają na stworzenie prezentacji interaktywnych, rozszerzenie funkcjonalności o testowanie wiedzy ucznia oraz wybór prezentowanych materiałów w oparciu o jego obecną wiedzę czy wprowadzenie elementów symulacji w celu wyrobienia umiejętności zastosowania zdobytej wiedzy teoretycznej w praktyce.

Takie właśnie publikacje elektroniczne, rozszerzone o możliwości udostępniane przez nowoczesne technologie zapisu, prezentacji i automatyzacji stanowią obecnie podstawę opracowania materiałów dydaktycznych w nauczaniu zdalnym, szczególnie w przypadku nauczania w trybie asynchronicznym.

Rzeczą na którą należy zwracać baczną uwagę jest zgodność narzędzia służącego tworzeniu treści z ustalonymi standardami e-learningu, co pozwoli wykorzystać stworzone elementy kursów jako część większego systemu zdalnego nauczania lub bazy wiedzy. Przy braku zgodności ze standardami użyteczność tego typu narzędzia będzie bardzo ograni-

czona. Niestety, z tego właśnie powodu, duża część dostępnych obecnie produktów tego typu, może służyć jedynie do tworzenia prezentacji będących zamkniętymi całościami, bez możliwości powiązania ich w całościowy kurs lub część zaawansowanego systemu edukacji zdalnej.

W dalszej perspektywie przewiduje się oderwanie narzędzi służących tworzeniu zawartości merytorycznej kursów od internetowych technologii służących **opisowi oraz prezentacji treści w ostatecznej formie** przedstawianej użytkownikowi. Powodem tego będzie przeniesienie roli tworzenia gotowej do użycia przez ucznia formy kursu na barki zaawansowanego systemu elektronicznego nauczania i zarządzania treścią, podczas gdy narzędzia tworzenia kursów generować będą tylko określoną bazę wiedzy i wytyczne dotyczące jej możliwego wykorzystania. Jednak i tu technologie internetowe przychodzą nam z pomocą – a mianowicie technologie zapisu i zarządzania informacją, których przykładem i podstawą może być język XML.

2.2. Produkty służące transmisji obrazu i dźwięku

Częstą praktyką w przypadku zdalnego nauczania są próby bezpośredniego użycia metod nauczania klasycznego przy wykorzystaniu nowego medium. Najprostszą metodą realizacji takiego podejścia wydaje się zastosowanie systemu telekonferencyjnego, w celu stworzenia grupy osób mogących porozumiewać się za pomocą fonii jak i wizji, uczestnicząc w zajęciach prowadzonych w trybie synchronicznym.

Niestety rozwiązanie takie posiada wiele wad. Pod względem technicznym jest zarówno trudne w realizacji jak i kosztowne. Ilość przesyłanych siecią komputerową danych jest niewspółmiernie duża do przekazywanych uczniom wiadomości, a wymagania stawiane wyposażeniu użytkowników – wysokie. Dodatkowo, generowany ruch wymaga wysokiej i stałej, a zatem trudnej do zapewnienia, jakości obsługi.

Z pomocą przychodzą tu, opracowane dla potrzeb sieci Internet, technologie kompresji i transmisji zawartości audio/video (*Streaming Media Services*). Jednakże, nawet pomimo ich zastosowania, koszty takiego systemu rosną znacząco wraz ze wzrostem liczby jego użytkowników, co jest wyjątkiem w nowoczesnych systemach zdalnej edukacji i limituje przydatność takiego rozwiązania.

W praktyce okazuje się często, że metody dydaktyczne sprawdzające się w nauczaniu klasycznym (jak na przykład: sposób prowadzenia wykładów czy ćwiczeń), wymagają przynajmniej znaczącego dostosowania do nowego medium, a zatem nieprawdziwy jest argument o łatwości wprowadzania zdalnej edukacji tą metodą.

Nie należy jednak zapominać iż zastosowanie wideokonferencji jest w dalszym ciągu niezastąpionym sposobem na uzyskanie bezpośredniego kontaktu z nauczycielem, jednakże wykorzystanie ich jako podstawy systemu zdalnego nauczania nie jest, z pewnością, rozwiązaniem optymalnym.

2.3. Systemy zarządzania nauczaniem LMS

Opisane wyżej klasy produktów e-learningowych koncentrują się prawie wyłącznie na przekazywaniu treści dydaktycznej, pozostawiając na uboczu sprawy organizacji procesu nauczania zarówno dla każdego z uczniów, jak i dla wszystkich uczestników zdalnej nauki, jako całości.

Organizacją procesu zdalnego nauczania zajmują się systemy klasy LMS (*Learning Management System* – system zarządzania nauczaniem). LMS to system komputerowy,

który automatyzuje proces zarządzania, administracji, śledzenia i raportowania wszelkich działań związanych ze szkoleniami w ramach jednej lub nawet wielu instytucji. Systemy informatyczne tego typu umożliwiają tworzenie, przechowywanie, zarządzanie i publikowanie treści dydaktycznej w postaci kursów multimedialnych, stron WWW, a także organizowanie procesu dydaktycznego, kontrolowanie postępów ucznia, zapisywanie i kontrolowanie dostępu do kursów oraz raportowanie. System typu LMS można podzielić na 4 podstawowe części:

Moduł zarządzania szkoleniami

Moduł ten oferuje funkcje zarządzania procesem zdalnego nauczania, na przykład: tworzenie harmonogramów zajęć i kursów, budowę katalogu dostępnych zasobów, import i udostępnianie kursów osobom szkolonym, zarządzanie zasobami takimi jak sale wykładowe, rozliczanie i śledzenie opłat za kursy...

Moduł zarządzania szkoleniami obsługuje też funkcje śledzące proces nauki (które, na przykład: zapisują dane dotyczące ścieżki nauczania dla każdej osoby szkolonej z osobna, monitorują postępy kształconych osób, ewidencjonują koszty procesu nauki), oraz raportujące, które oferują różnego rodzaju raporty dotyczące rezultatów nauki, satysfakcji studentów, ich zachowań w procesie edukacyjnym oraz inne raporty o charakterze administracyjnym.

Moduł zdalnego samokształcenia

Moduł zdalnego samokształcenia pozwala osobom szkolonym na korzystanie ze zdalnych kursów. Jest to spersonalizowana witryna udostępniająca materiały szkoleniowe przeznaczone dla danej osoby.

Moduł budowy kursów

Moduł budowy kursów udostępniany jest prowadzącemu szkolenie. Umożliwia on zwykle tworzenie szkieletu kursów, wypełnianie ich treścią tekstową oraz multimedialną, budowę szablonów ćwiczeń oraz wypełnianie ich konkretnymi zadaniami a także inne, proste lub średnio-zaawansowane działania związane z materiałem szkoleniowym.

Moduł komunikacyjny

Moduł komunikacyjny pozwala na komunikację pomiędzy osobami uczestniczącymi w szkoleniu lub pomiędzy nimi a prowadzącym. Oferuje on zazwyczaj dobrze znane użytkownikom Internetu metody kontaktu w sposób:

1. synchroniczny, jak na przykład: chat, współdzielenie ekranu, przekaz głosowy czy webcasting,
2. asynchroniczny, jak email czy listy dyskusyjne.

2.4. Systemy zarządzania treścią nauczania LCMS

O ile systemy typu LMS koncentrują się przede wszystkim na organizacji kursu zdalnego nauczania, to systemy LCMS (*Learning Content Management System* – system zarządzania treścią nauczania) służą projektowaniu, tworzeniu, składowaniu i dostarczaniu spersonalizowanych materiałów szkoleniowych w postaci *obiektów szkoleniowych* (*learning objects*). Poza tworzeniem, przechowywaniem, zarządzaniem i publikowaniem treści dydaktycznej w postaci kursów multimedialnych, stron WWW, pozwalają dodatkowo, na organizowanie procesu dydaktycznego, kontrolowanie postępów ucznia, zapisywanie i kontrolowanie dostępu do kursów oraz raportowanie.

Podstawowe elementy składowe systemu LCMS to:

Moduł składnicy obiektów

Moduł składnicy obiektów obsługuje bazę danych, zawierającą wszelkie elementy wchodzące w skład kursów obsługiwanych przez system. Z tej bazy, obiekty tworzące szkolenie przesyłane są do jego uczestników. Dzięki wykorzystaniu informacji zawartych w repozytorium, można z obiektów zawartych w bazie generować kursy w różnej (zależnej od potrzeb) postaci, na przykład: strony Web odpowiedniej dla łącz różnej przepływności, kursu dystrybuowanego na płytach CD lub DVD, materiałów przeznaczonych do druku, stron w formacie odpowiednim dla urządzeń PDA itp.

Moduł tworzenia kursów

Moduł ten odpowiada za tworzenie obiektów składowych kursów, tzw. SCO (*Sharable Content Objects*). Umożliwia zarówno budowę nowych obiektów, jak i wykorzystanie gotowych szablonów oraz rozbudowane funkcje zarządzania obiektami już istniejącymi (kopiowanie, kasowanie, edycję...).

Dysponując już obiektami SCO można, za pomocą tego modułu, zbudować z nich kompletny kurs. Ponadto, ponieważ tworzenie zawartości merytorycznej kursu jest najczęściej działaniem zespołowym, systemy klasy LCMS umożliwiają również zaawansowane zarządzanie pracą grupową nad kursami.

Moduł dostarczania kursów

Moduł dostarczania kursów pozwala udostępniać kursy studentom według ustalonych profili. Umożliwia również śledzenie postępów osoby szkolonej oraz raportuje wyniki ćwiczeń, testów, pytań, itp.

Moduł administracyjny

Moduł administracyjny oferuje możliwości odpowiadające obszarowi zainteresowań systemu LMS, czyli umożliwia zarządzanie procesem nauki: zarządza kontami osób szkolonych, mechanizmami udostępniania im kursów, śledzi postępy w nauce oraz prowadzi inne czynności administracyjne. Moduł ten może zostać zintegrowany z pełnym systemem klasy LMS, zapewniającym bardziej zaawansowaną funkcjonalność.

2.5. Korporacyjne platformy edukacji zdalnej

Środowisko korporacyjne jest szczególnie interesujące w przypadku udostępniania usług elektronicznych. Mamy tu do czynienia z zamkniętą grupą użytkowników, o których dysponujemy znaczną wiedzą i którym udostępniamy znaczną liczbę usług. Środowisko to jest zatem idealnym miejscem do wdrażania systemów o znacznym stopniu automatyzacji, umożliwia zastosowanie najbardziej rozbudowanych systemów zdalnego nauczania, oferujących pełną kontrolę nad procesem dydaktycznym. Takie rozwiązania okażą się w tym przypadku najbardziej efektywne i ekonomiczne.

Ze względu na konieczność stałego aktualizowania i podnoszenia wiedzy i umiejętności pracowników, duże korporacje są również jedną z najbardziej zainteresowanych rozwojem e-learningu stron.

3. PRZEGLĄD TECHNOLOGII INTERNETOWYCH

W początkowych latach rozwoju nauczania wspomagane komputerowo (CBI – *Computer Based Instruction*) wielu specjalistów niezależnie pracowało nad pokonaniem

najbardziej ewidentnej przeszkody, stojącej na drodze do jego popularyzacji i szerokiego wykorzystania – a mianowicie trudności w tworzeniu i w konsekwencji braku materiałów dydaktycznych w formie elektronicznej. Dodatkową trudnością był tu fakt, iż specjaliści z dziedzin innych niż informatyka nie posiadali na ogół wystarczającej wiedzy na temat obsługi sprzętu i oprogramowania, aby tworzyć komputerowo wspomagane kursy ze swoich dziedzin.

Prace te zaowocowały powstaniem wielu niezależnych technologii opisu treści dydaktycznej oraz narzędzi pozwalających na łatwe przetwarzanie zawartości dydaktycznej na formę elektronicznych kursów. Kursy te stanowiły jednakże zamknięte produkty, uruchamiane na komputerze użytkownika, o ściśle określonej zawartości i strukturze, a w dodatku korzystające z technologii charakterystycznych dla ich producenta czy narzędzia którym je stworzono.

Popularyzacja Internetu i gwałtowny rozwój towarzyszących mu technologii i standardów, skierował główny nurt prac nad systemami zdalnego nauczania na zupełnie inne tory. Kursy CBI przestały mieć postać pakietów przeznaczonych do uruchomienia na komputerze użytkownika końcowego (client-based), a zostały przystosowane do dostarczania on-line, korzystając z gotowych standardów i technologii zapisu i prezentacji istniejących w sieci Internet.

Ta zmiana, z architektury typu client-based, na architekturę web-based thin-client, umożliwiła powstanie **zintegrowanych systemów** zdalnego nauczania, w miejscu izolowanych kursów, oraz ogromnie obniżyła koszty ich wdrażania.

W chwili obecnej główny nurt prac nad systemami e-learningu skierowany jest na opracowanie metod sterowania dostarczaniem zawartości dydaktycznej, a nie samego jej tworzenia i prezentacji. Jak wynika z powyższego podejścia, systemy elektronicznego nauczania muszą mieć możliwość pracy nad zawartością dydaktyczną utworzoną za pomocą różnych, szybko rozwijających się technologii internetowych, gdyż to właśnie na nich polegamy w kwestiach zapisu i prezentacji zawartości.

Ta radykalna zmiana podejścia do nauczania elektronicznego i kierunku jego rozwoju została podkreślona również zmianą nazwy tego rodzaju aktywności z nauczania wspomaganego komputerowo (CBI), na nauczanie zarządzane komputerowo (*Computer Managed Instruction* – CMI).

Wprowadzanie wyraźnego podziału między technologiami tworzenia, opisu i prezentacji treści dydaktycznych, a technologiami kontroli i kierowania samym procesem kształcenia pozwoliło skoncentrować się na tworzeniu systemów ITS podnoszących efektywność elektronicznego nauczania, a odpowiedzialność za zapewnienie możliwości wykorzystania nowych mediów w procesie dydaktycznym złożyć na barki technologii internetowych i ich niezależnego rozwoju.

3.1. Technologie tworzenia i prezentacji dokumentów

Technologie tworzenia i prezentacji dokumentów pełnią dziś główną rolę w budowie kursów zdalnego nauczania, a wszystko wskazuje, że ich znaczenie również w przyszłości nie ulegnie zmniejszeniu. Zmieni się natomiast z całą pewnością ich rola.

Obecnie materiał dydaktyczny jest tworzony przez autora w formie gotowej do przedstawienia użytkownikowi, niezależnie od stosowanej przez niego platformy sprzętowo/programowej. W takiej formie też przechowywany jest przez system e-learningu.

W przyszłości technologie służące opisowi ostatecznego wyglądu i funkcjonowania kursu, używane będą dopiero w ostatniej fazie dostarczania użytkownikowi treści kursu.

Sama baza materiałów dydaktycznych zawierać będzie obiekty opisane za pomocą języka o wyższym stopniu abstrakcji i dopiero na ich podstawie system wygeneruje (z pomocą opisywanych technologii) dokumenty przeznaczone dla końcowego użytkownika.

Technologie opisu i prezentacji dokumentów wykorzystywane w Internecie pozwalają na opisanie praktycznie dowolnego dokumentu, a co równie ważne, mogą być serwowane z szerokiej gamy mediów – od serwera umieszczonego na innym kontynencie, do płyty DVD używanej lokalnie.

Najczęściej wykorzystywane są technologie pozwalające na tworzenie statycznych dokumentów:

- HTML, pozwalający na łatwe przenoszenie tekstu na różne platformy sprzętowe i programowe,
- języków opisu animacji/3D itp. (FLASH, Shockwave, WMRL),
- metody kompresji i zapisu obrazu statycznego (JPEG, PNG, GIF),
- metody kompresji i zapisu obrazu ruchomego (RealVideo, ASF, AVI),
- metody kompresji i zapisu dźwięku (MPEG Layer-3, RealAudio).

Gwałtowny rozwój przeżywają również technologie pozwalające na tworzenie dokumentów dynamicznych, których treść nie jest stała, lecz zależy od różnych czynników lub zmienia się w reakcji na działania użytkownika. Najpopularniejsze technologie tego typu to:

- języki skryptowe wykonywane po stronie klienta (JavaScript, VBScript),
- języków skryptowych wykonywanych po stronie serwera (PHP, Active Server Pages),
- narzędzi tworzenia oprogramowania niezależnego od platformy sprzętowej (JAVA, ActiveX).

Każda z tych technologii jest już standardem, obsługiwanym przez większość systemów operacyjnych, wspieranych przez mechanizmy automatycznej aktualizacji. Powoduje to, że możliwa jest z ich pomocą nieomal idealna realizacja idei „thin-client”, gdzie użytkownik jest w stanie skorzystać z praktycznie dowolnej usługi, bez konieczności instalowania dodatkowego oprogramowania poza systemem operacyjnym.

3.2. Technologie przetwarzania rozproszonego i sieciowe

Chociaż zwykle mówiąc o technologiach internetowych, mamy na myśli standardy opisu treści oraz jej późniejszej prezentacji, nie jest to jedyna dziedzina, której one dotyczą. Równie ważne, choć nie tak oczywiste dla przeciętnego użytkownika, jest fizyczne przekazywanie danych pomiędzy węzłami sieci oraz zapewnienie programistom aplikacji przeznaczonych do pracy w Internecie, narzędzi jak najefektywniej wykorzystujących zalety tego środowiska.

Narzędzia programistyczne, takie jak na przykład.: Microsoft Visual Studio .NET, czy większość środowisk typu RAD (Rapid Application Developement), zawierają szereg funkcji ułatwiających wykorzystanie środowiska sieci Internet do tworzenia systemów rozproszonych:

- XML (*Extensible Markup Language*) – uniwersalny język opisu dokumentów i przekazywania informacji w zestandaryzowanej formie. Stanowi obecnie podstawę wielu innych technologii, umożliwiając przenoszenie informacji w formie łatwej w interpretacji dla systemów informatycznych. Stanowi uproszczoną i dostosowaną do środowiska Internetu wersję języka SGML (ISO 8879:1985).

- SOAP (*Simple Object Access Protocol*) – prosty, wykorzystaniu XML oraz innych, gotowych protokołów wysokich warstw, sposób zdalnego wywoływania obiektów lub procedur w środowisku rozproszonym. Umożliwia procesom łatwą komunikację niezależnie od środowiska ich działania czy też faktycznie wykorzystanego do komunikacji protokołu (HTTP, SMTP...).
- CORBA (*Common Object Request Broker Architecture*) – to system zdalnego, niezależnego od platformy sprzętowej, systemu operacyjnego i języka programowania, dostępu do obiektów. Bardziej skomplikowany od SOAP lecz oferujący większe możliwości.
- UDDI (*Universal Description, Discovery and Integration Service*) – usługa udostępniająca klientom mechanizmy wyszukiwania innych usług Web. UDDI stanowi interfejs umożliwiający dynamiczne połączenie się z usługą udostępnianą przez innego usługodawcę.

Obecność tego rodzaju mechanizmów umożliwia pisanie aplikacji w sposób pozwalający im na współpracę niezależnie od fizycznej lokalizacji i rodzaju urządzeń na których pracują, jak długo zachowana zostaje komunikacja sieciowa.

Właściwość tego typu okaże się niezastąpiona w budowie nowoczesnych systemów zdalnej edukacji, szczególnie z użyciem opisanego dalej standardu SCORM, który kładzie nacisk na łatwość dostępu do informacji zmagazynowanej w wielu lokalizacjach.

Jak już wspomniano, do poprawnego funkcjonowania którejkolwiek z opisanych powyżej technologii, konieczne jest zachowanie komunikacji sieciowej – a dokładniej, konieczne jest zapewnienie komunikacji o jakości i parametrach odpowiednich do potrzeb danej usługi.

O ile rozwój technologii sieciowych pozwala na transmisję danych z prędkościami adekwatnymi do aktualnych potrzeb, to specyfika pakietowej sieci IP nie zapewnia jakiegokolwiek, ściśle określonej jakości obsługi w sieci. Prace nad rozwiązaniem tego problemu trwają już od dawna i zaowocowały powstaniem pewnych rozwiązań (jak przykładowo: protokół rezerwacji zasobów sieciowych RSVP i różne metody kształtowania ruchu), lecz wciąż w ogólnym przypadku, w sieci Internet nie możemy liczyć na stałe parametry obsługi ruchu sieciowego.

4. PRACE STANDARYZACYJNE

Oczywiste zalety zdalnego nauczania z wykorzystaniem mediów elektronicznych doprowadziły do konieczności doprowadzenia e-learningu do poziomu w pełni zestandaryzowanej technologii.

Dzięki współpracy wielu organizacji, takich jak: Advanced Distributed Learning (ADL), Alliance of Remote Instructional Authoring & Distribution Networks for Europe (ARIADNE), Aviation Industry CBT Committee (AICC), IMS Global Learning Consortium, Inc. the (IMS), the Institute of Electrical and Electronics Engineers (IEEE), stworzono standard SCORM (**S**harable **C**ontent **O**bject **R**eference **M**odel) [5-8, 10]. Jego zadaniem jest zdefiniowanie technicznych podstaw systemu zdalnego nauczania opartego na technologiach internetowych. Jest zbiorem zasad definiujących sposoby i formaty opisu, przechowywania i dostarczania użytkownikowi treści dydaktycznych. Zakłada on użycie tzw. SCO (*Sharable Content Object*) – obiektów zawierających treść dydaktyczną i zbudowanie z nich rozproszonej bazy wiedzy, którą można następnie wykorzystywać w celu tworzenia kursów zdalnego nauczania, systemów eksperckich itp.

Cechą charakterystyczną standardu SCORM, jest odwołanie się do istniejących technologii w celu realizacji określonych zadań. Tworzy on model odwołujący się do istniejących standardów, procedur i specyfikacji w celu realizacji określonych zadań wysokiego poziomu.

Zaplanowano realizację następujących postulatów:

- accessibility – możliwość wyszukania i dostępu do treści z odległej lokalizacji i dostarczenia jej do wielu punktów docelowych,
- adaptability – możliwość dostosowania nauczania dla potrzeb jednostki i organizacji,
- affordability – zmniejszenie kosztów i zwiększenie efektywności e-learningu,
- durability – brak konieczności wprowadzania daleko idących zmian w celu dostosowania systemu do zachodzącego postępu technologicznego,
- interoperability – uniezależnienie zawartości dydaktycznej od sprzętowej i programowej platformy e-learningu oraz narzędzi tworzenia treści,
- reusability – możliwość użycia komponentów istniejących kursów w nowych zastosowaniach i kontekstach,

Dodatkowym założeniem standardu SCORM, jest uznanie wykorzystania technologii internetowych za najlepszy sposób maksymalizacji dostępu do systemów elektronicznej edukacji oraz zapewniania realizacji wymienionych powyżej postulatów.

5. ZAKOŃCZENIE

Technologie internetowe stanowią jeden z dwóch podstawowych filarów nowoczesnych systemów zdalnej edukacji – odpowiadają za opis, dostarczanie i prezentację treści dydaktycznej. Możliwość przesunięcia ciężaru realizacji tych elementów nauczania na barki niezależnego rozwoju technologii internetowych, pozwoli projektantom systemów CMI na skoncentrowaniu uwagi na sterowaniu samymi procesami nauczania. Spowoduje to znaczący rozwój, zwiększenie efektywności i popularyzację systemów ITS, które mają szansę stać się najefektywniejszą metodą kształcenia.

Warto zauważyć, iż zawartość dydaktyczna opracowana ze wsparciem technologii internetowych, nadaje się do prezentacji w dowolny praktycznie sposób i za pomocą niemal dowolnego medium (sieci komputerowe przewodowe i bezprzewodowe, CD-ROM, DVD, czy nawet wydruk), co w połączeniu z możliwościami oferowanymi przez technologie sieciowe oraz Internet pozwala na powszechny i łatwy dostęp do tak realizowanego kształcenia. Powyższe cechy, w połączeniu z szacowaną dużą efektywnością systemów ITS, przewyższającą efektywność większości stosowanych obecnie, klasycznych, metod dydaktycznych, a zbliżającą się do pracy z prywatnym nauczycielem, gwarantują dalszy rozwój systemów CMI.

Należy sobie ponadto uświadomić, że tak wysoka efektywność w połączeniu z opisanymi w artykule cechami technologii internetowych prowadzi do dezaktualizacji terminu „system zdalnego nauczania” – systemy tego typu opłaci się zastosować w przypadku kursów o każdej formule (w tym stacjonarnych), a cechy technologii internetowych uniezależnią dostęp do nich od wzajemnej lokalizacji studenta i systemu.

BIBLIOGRAFIA

- [1] Gettinger M.: *Individual differences in time needed for learning: A review of the literature*, 1984.
- [2] Graesser A.C., Person N.K.: *Question asking during tutoring*, American Educational Research Journal, 1994.
- [3] Grabowska A.: *Technologie internetowe w kształceniu ustawicznym inżynierów*, Magazyn Edukacji Wirtualnej MEWa, 2002.
- [4] Chmielewski A.: *Nowoczesne rozwiązania technologiczne w nauczaniu przez Internet na przykładzie systemów klasy LCMS*, 2002.
- [5] Advanced Distributed Learning, "SCORM 2004 Overview", 30 January 2004, <http://www.adlnet.org>
- [6] Advanced Distributed Learning, "SCORM 2004 Content Aggregation Model", 30 January 2004, <http://www.adlnet.org>
- [7] Advanced Distributed Learning, "SCORM 2004 Runtime Environment", 30 January 2004, <http://www.adlnet.org>
- [8] Advanced Distributed Learning, "SCORM 2004 Sequencing & Navigation", 30 January 2004, <http://www.adlnet.org>
- [9] AICC/CMI "CM1001 Guidelines for Interoperability, Version 3.4.", 23 October 2000, <http://www.aicc.org/>
- [10] Learning Systems Architecture Lab, "Technical Evolution of SCORM", 18 August 2003, <http://www.lsal.cmu.edu>
- [11] World Wide Web Consortium, Internet Technologies Whitepapers, January 2004, <http://www.w3.org>
- [12] U.S. Department of Education, "The Power of the Internet and Learning: Moving from Promise to Practice, Report of the Web-Based Education Commission", 15 December 2000.

INTERNET TECHNOLOGIES IN REMOTE LEARNING SYSTEMS

Summary

Remote learning services had become more and more popular nowadays. Not only private customers, but also corporate employers, are interested in new learning methods, which promise to minimize effect of traditional barriers in access to education conducted by classical methods.

Most of today's remote learning systems use web-technologies for many aspects of creation, storage and serving a didactic content. Their use in such systems focuses on widely understood information management, including advanced mechanisms for input, storage, indexing and searching etc. As internet technologies are now used as sole solution for this purpose in modern, computer based remote learning, their impact on development and advancement of e-learning systems is paramount.

This article presents main types of remote learning products and systems and their basic classification in light of current standardization process (SCORM standard), followed by description and classification of internet technologies most often used in such systems, as well as analysis of their impact on design of remote learning systems and efficiency of didactic process.

Anna Grabowska, Jerzy Durslewicz, Bartłomiej Antczak, Tomasz Skwarek

**Centrum Edukacji Niestacjonarnej, Politechnika Gdańska
WiedzaNet Sp. z o.o., Incenti S.A., Oracle**

MODEL SZKOLENIA "BLENDED LEARNING" Z WYKORZYSTANIEM PLATFORMY ORACLE I-LEARNING

Streszczenie

W artykule zaproponowano modele organizacyjne szkoleń „blended learning”, które pokazują możliwości współpracy firm prywatnych z instytucjami edukacyjnymi w dziedzinie e-learningu. W ramach wspólnego eksperymentu firm Oracle [1], [3], Incenti S.A. [4], WiedzaNet Sp. z o.o. [5] oraz Centrum Edukacji Niestacjonarnej Politechniki Gdańskiej w semestrze letnim roku akademickiego 2003/2004 udostępniony będzie kurs dla studentów Wydziału Inżynierii Lądowej z wykorzystaniem hostowanej platformy i-learning firmy Oracle.

1. WSTĘP, CZYLI CZYM JEST „BLENDED LEARNING”

Angielski termin “blended learning” [2] oznacza kształcenie z wykorzystaniem różnych metod i mediów. Najogólniej ten rodzaj kształcenia wymaga odpowiednio przygotowanych obiektów wiedzy, które są przekazywane za pomocą właściwie wybranych technologii uwzględniających preferencje uczącego się (czas, miejsce, metody), tak aby w konsekwencji przekazać wiedzę (umiejętności) stosując strategię „to czego potrzebuję się nauczyć w danym momencie” (ang. *just-what-I-need, just-in-time*).

„Blended learning” traktowano do niedawna jako połączenie tradycyjnej nauki w klasie z elementami e-learningu. W najprostszym rozumieniu interpretowano tę formę kształcenia jako połączenie zajęć w klasie z nauką przez Internet lub w intranecie.

W chwili obecnej można wyróżnić następujące metody wspierające formułę „blended”:

Metody synchroniczne wymagające fizycznej obecności w danym miejscu i czasie (wykłady, laboratoria, warsztaty organizowane w salach w określonych godzinach, pod kierunkiem obecnego w klasie nauczyciela).

Metody synchroniczne wykorzystujące środowiska wirtualne (klasy wirtualne, internetowe transmisje audio, wideo, seminaria online, czat, czyli rozmowy online).

Metody asynchroniczne (dokumenty, materiały w formacie WWW, elektroniczne moduły szkoleniowe zawierające testy samosprawdzające, symulacje, materiały audio, wideo, fora dyskusyjne).

Aby przybliżyć koncepcję metody „blended learning” zaproponowano następujący scenariusz [1]:

- Uruchomienie portalu informacyjnego dot. szkolenia, który zawiera podstawowe informacje ułatwiające właściwe wykorzystanie zasobów edukacyjnych.
- Spotkanie w klasie (ang. „kick-off”) w celu przedstawienia uczestników i instruktorów oraz stworzenia zespołów, których członkowie będą zaangażowani we wspólne działania.
- Zajęcia w trybie asynchronicznym, wykorzystujące szkolenia elektroniczne.
- Dyskusje w trybie asynchronicznym pozwalające uczestnikom szkoleń na wymianę doświadczeń oraz prezentację własnych prac.
- Seria spotkań z instruktorami w trybie synchronicznym (w klasie lub w formie rozmów online).
- Test samosprawdzający określający poziom zdobytej wiedzy (umiejętności) ewentualnie egzamin certyfikujący.
- Ankieta, oceniająca szkolenie, której wyniki pozwolą na modyfikację następnych kursów.

Zanim uruchomione zostanie nowe szkolenie w trybie „blended learning” należy zwrócić uwagę na następujące elementy procesu kształcenia [1]:

- Grupa celowa w kontekście posiadanego poziomu wiedzy z zakresu oferowanego szkolenia, preferencji dotyczących zakresu szkolenia, preferowanych stylów nauki, lokalizacji miejsca spotkań, motywacji.
- Zawartość merytoryczna szkolenia z punktu widzenia optymalnego formatu przekazu wiedzy.
- Koszty uruchomienia szkolenia powinny uwzględniać koszt przygotowania materiałów oraz ich dystrybucji. Interaktywne, multimedialne szkolenia są ekonomicznie uzasadnione w przypadku dużej liczby uczestników, co najmniej kilka tysięcy.
- Infrastruktura może w znaczącym stopniu decydować o sposobie realizacji szkolenia. Należy wziąć pod uwagę dostęp do nowych technologii (Internet, urządzenia mobilne, szybkość łączny).

2. MODELE WSPÓŁPRACY INSTYTUCJI EDUKACYJNEJ Z FIRMĄ KOMERCYJNĄ W ZAKRESIE I-LERNINGU

Założono, że w zaproponowanym modelu współpracy instytucji edukacyjnej, zwanej dalej centrum edukacyjnym z firmami komercyjnymi w zakresie wdrażania specjalistycznych szkoleń elektronicznych wykorzystana będzie metoda „blended learning”.

Na wstępie zdefiniowano następujące elementy modelu:

Platforma – środowisko czyli zestaw narzędzi i technologii, które mogą zostać wykorzystane do stworzenia i udostępniania materiałów oraz komunikacji wymaganej do przeprowadzenia kursu. Wyróżnia się dwa rodzaje platform LMS i LCMS. System LMS (*Learning Management System*) to system informatyczny dostarczający środowisko do zarządzania procesami kształcenia w firmach i instytucjach oraz oprogramowanie, które pozwala wypełniać to środowisko kursami i efektywnie nimi zarządzać. Zarządzanie procesem kształcenia powinno umożliwiać organizację struktur szkoleniowych, takich jak grupy uczestników szkoleń, grupy kursów, administrowanie użytkownikami i ich prawami dostępu do poszczególnych funkcji i zasobów systemu. Zwykle dostęp do kursów jest możliwy przez standardowe przeglądarki internetowe, dzięki czemu każdy uczestnik

szkolenia bez względu na swoje miejsce pobytu (biuro, dom) może uczestniczyć w kursie. Typowe funkcje systemu LMS to m.in. zarządzanie użytkownikami, importowanie i odtwarzanie kursów różnych dostawców, zarządzanie kursami, śledzenie i weryfikacja postępów użytkowników, raportowanie wyników, udostępnianie narzędzi komunikacyjnych takich jak e-mail, forum dyskusyjne, czat. System LCMS (*Learning Content Management System*) to system informatyczny umożliwiający tworzenie i publikowanie kursów, przy czym większość systemów LCMS ma również wbudowaną funkcjonalność systemów LMS.

Kurs – wszystkie materiały tak tekstowe jak i multimedialne, oraz inne elementy kursu Internetowego, z reguły opracowane już z wykorzystaniem Platformy.

Nauczyciel – w większości poważnych kursów element nieodzowny ciągu trwania szkolenia, koordynator i konsultant. Osoba taka musi posiadać odpowiednią wiedzę merytoryczną z zakresu prowadzonego kursu, ale musi także być obeznana z metodami i specyfiką szkolenia na odległość. Należy jednak zaznaczyć że istnieje pewna grupa kursów, często określanych mianem „stand alone”, które nie wymagają nadzoru nauczyciela.

Egzamin – jedyny wiarygodny sposób weryfikacji zdobytej wiedzy i umiejętności. Jest to także jedyny element kursu jaki nie może być prowadzony na odległość ze względu na konieczność upewnienia się co do rzeczywistych umiejętności egzaminowanego. Do przeprowadzenia egzaminu wymagane są odpowiednie warunki lokalowo sprzętowe.

Egzaminator – osoba nadzorująca przebieg egzaminu.

Certyfikat – można wyróżnić 2 rodzaje certyfikatów. Po zakończeniu kursu każdy z uczestników może otrzymać certyfikat odbycia kursu, jednakże tylko ci, którzy pomyślnie zdadzą egzamin mogą otrzymać certyfikat poświadczający zdobycie określonych umiejętności. Certyfikat może być wydawany przez firmę komercyjną, centrum edukacyjne lub przez oba podmioty wspólnie.

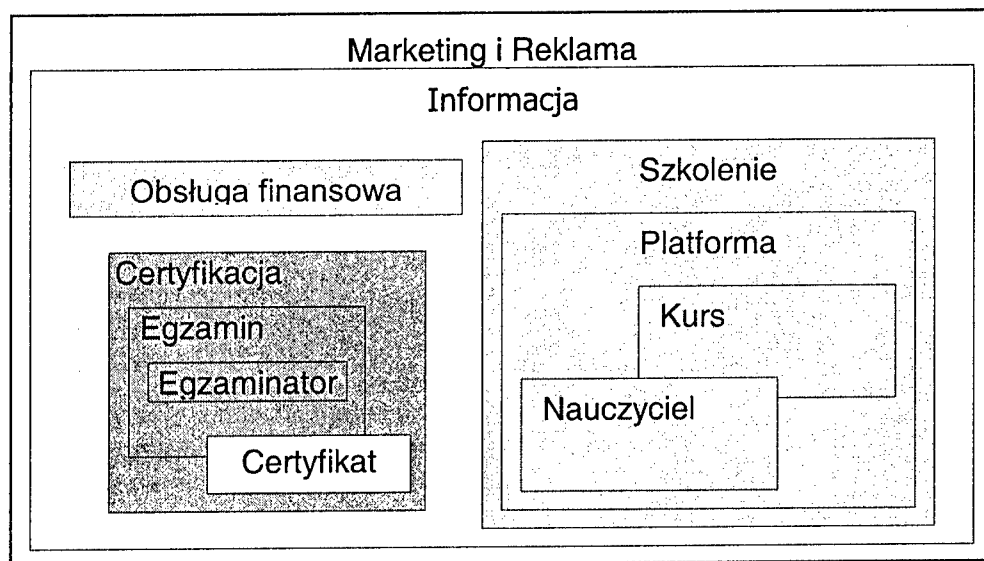
Marketing/Reklama – wszelkie działania mające na celu zdobycie klientów.

Punkt Informacyjny – miejsce (rzeczywiste lub wirtualne), w którym potencjalny klient może zasięgnąć informacji na temat kursów egzaminów itp.

Obsługa finansowa – punkt obsługi finansowej, z reguły zależny od przyjętego modelu współpracy, choć najczęściej ograniczający się do numeru odpowiedniego konta bankowego.

Model 1: Centrum Certyfikacji

Zorganizowanie certyfikacji stanowi bardzo często duży problem dla komercyjnej firmy, ze względu na duże koszty jakie trzeba ponieść, aby uruchomić centrum certyfikacji. Wiąże się to z wynajęciem lokalu, zakupem komputerów i oprogramowania oraz przygotowaniem i zatrudnieniem kadry. Wszystko to z reguły jest już dostępne w jednostce edukacyjnej, dzięki czemu bardzo łatwo i szybko można podjąć współpracę w ramach certyfikacji. Sytuacja taka ma jeszcze jedną niepowtarzalną zaletę – firma komercyjna może, poprzez tak zrealizowaną sieć partnerów, niezwykle ułatwić życie swoim potencjalnym klientom, gdyż nie będą oni musieli pokonywać kilkuset kilometrów celem dostania się do centrum certyfikacji prowadzonego przez firmę, a tylko dużo krótszą odległość, jeżeli certyfikacja dokonywać się będzie w centrum edukacyjnym znajdującym się na pobliskiej uczelni.



W tym modelu, firma komercyjna zapewnia wszystkie kolejne etapy od stworzenia do zakończenia kursu (Platforma, Kurs, Nauczyciel, Punkt Informacyjny) zaś centrum edukacyjne udostępnia swoje zasoby lokalowe, sprzętowe i ludzkie celem przeprowadzenia Certyfikacji oraz pełni rolę lokalnego Punktu Informacyjnego.

Marketing/Reklama – w zależności od ustaleń prowadzone może być przez firmę, centrum, obie instytucje, lub podmiot zewnętrzny.

Certyfikat może wystawić firma, centrum, oba podmioty lub certyfikat zewnętrzny jeżeli firma/centrum/kurs posiada odpowiednie uprawnienia. Dopuszczenie do egzaminu następuje po ukończeniu odpowiedniego kursu oraz uiszczeniu opłat za kurs i egzamin, udokumentowanych potwierdzeniem wpłaty.

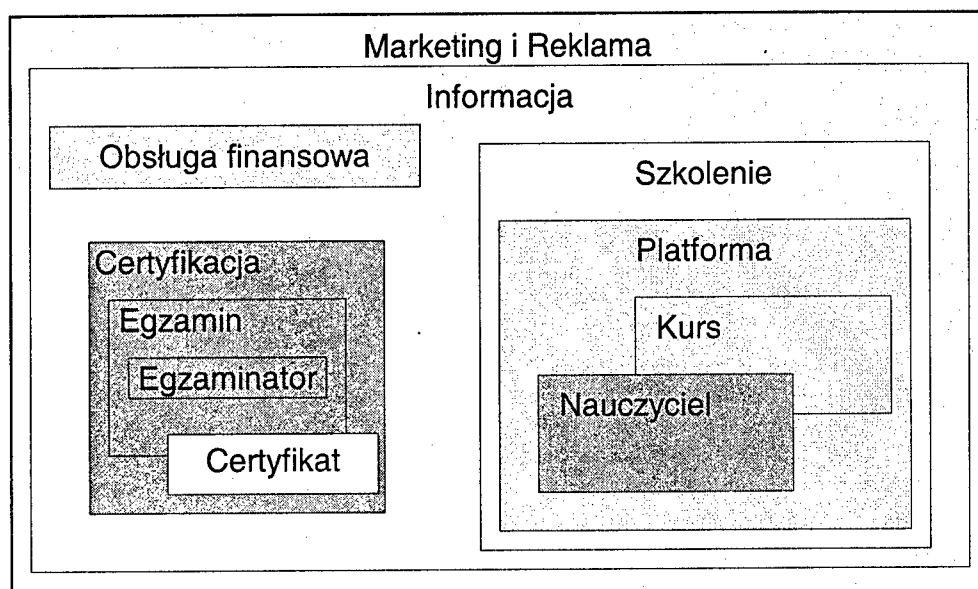
Obsługa finansowa prowadzona jest przez komercyjnego partnera, koszt certyfikacji jest wliczony w cenę kursu lub rozliczany całkowicie osobno. Opłaty wpływają na konto komercyjnego partnera, który następnie po zakończeniu Certyfikacji przelewa uzgodnioną kwotę na konto centrum edukacyjnego.

Model 2: Centrum Szkoleniowe

W wypadku drugiego modelu współpracy zdecydowanie rośnie znaczenie centrum edukacyjnego. Zaproponowano trzy warianty powyższego modelu.

Wariant 1:

W tym modelu centrum edukacyjne oferuje kurs partnera komercyjnego i jeżeli kurs tego wymaga, zapewnia kadrę do prowadzenia danego kursu. W wypadku centrów działających na uczelniach, forma ta jest szczególnie odpowiednia, gdy firma komercyjna chce trafić ze swoim kursem do studentów tej uczelni. Model ten może być szczególnie atrakcyjny, jeżeli kurs udostępniany będzie studentom bez opłat, np. w ramach zajęć, natomiast odpowiednie opłaty studenci będą ponosić, jeżeli zdecydują się na przystąpienie do egzaminu w celu uzyskania certyfikatu potwierdzającego zdobytą za pomocą kursu wiedzę.



Firma komercyjna: Platforma, Kurs, Punkt Informacyjny

Centrum edukacyjne: Nauczyciel, Egzamin, lokalny Punkt Informacyjny

Marketing/Reklama – w zależności od ustaleń prowadzone może być przez firmę, centrum, obie instytucje, lub podmiot zewnętrzny.

Certyfikat wystawiany przez firmę, centrum, oba podmioty lub certyfikat zewnętrzny jeżeli firma/centrum/kurs posiada odpowiednie uprawnienia. Dopuszczenie do egzaminu następuje po ukończeniu odpowiedniego kursu oraz uiszczeniu opłat za kurs i egzamin, udokumentowanych potwierdzeniem wpłaty.

Obsługa finansowa prowadzona jest przez komercyjnego partnera, koszt certyfikacji jest wliczony w cenę kursu lub rozliczany całkowicie osobno. Opłaty wpływają na konto komercyjnego partnera, który następnie po zakończeniu Certyfikacji przelewa uzgodnioną kwotę na konto centrum edukacyjnego.

Wariant 2:

W tym modelu centrum edukacyjne przeprowadza migrację kursu z własnej platformy lub przygotowuje nowy kurs do umieszczenia na platformie partnera. Jeżeli kurs tego wymaga, firma komercyjna zapewnia kadrę nauczycielską do prowadzenia danego kursu.

Firma komercyjna: Platforma, Nauczyciel, Punkt Informacyjny

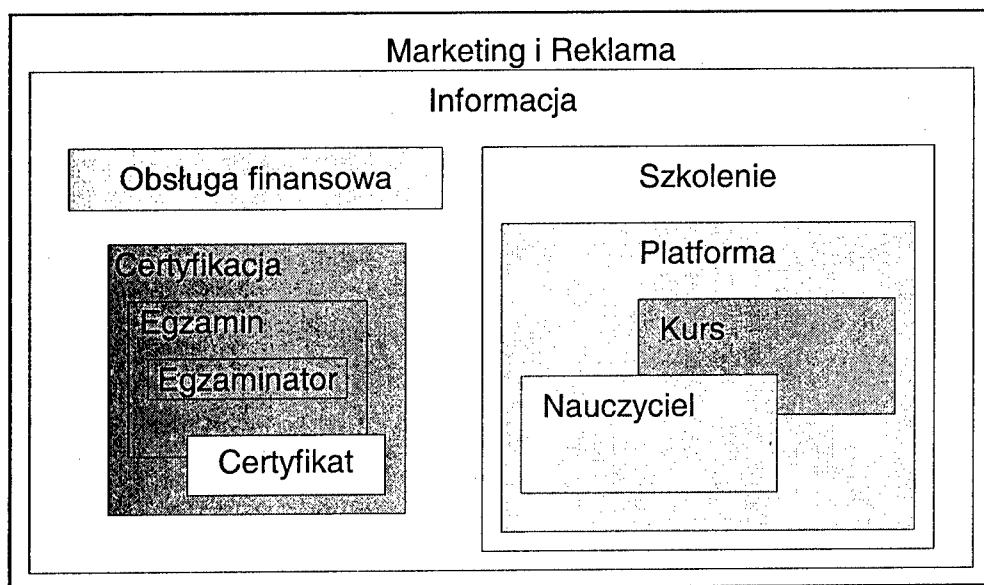
Centrum edukacyjne: Egzamin, Punkt Informacyjny, Kurs

Marketing/Reklama – w zależności od ustaleń prowadzone może być przez firmę, centrum, obie instytucje lub podmiot zewnętrzny.

Certyfikat wystawiany przez firmę, centrum, oba podmioty lub certyfikat zewnętrzny jeżeli firma/centrum/kurs posiada odpowiednie uprawnienia. Dopuszczenie do egzaminu następuje po ukończeniu odpowiedniego kursu oraz uiszczeniu opłat za kurs i egzamin, udokumentowanych potwierdzeniem wpłaty.

Obsługa finansowa prowadzona jest przez komercyjnego partnera, koszt certyfikacji jest wliczony w cenę kursu lub rozliczany całkowicie osobno. Opłaty wpływają na konto

komercyjnego partnera, który następnie po zakończeniu Certyfikacji przelewa uzgodnioną kwotę na konto centrum edukacyjnego



Wariant 3:

W tym modelu centrum edukacyjne przeprowadza migracje kursu z własnej platformy lub przygotowuje nowy kurs do umieszczenia na platformie partnera oraz, jeżeli kurs tego wymaga, zapewnia kadrę nauczycielską do prowadzenia danego kursu. Jest to najodpowiedniejszy model współpracy jeżeli partner komercyjny jest firmą, która funkcjonuje już na rynku szkoleń i chce rozszerzyć swoją aktualną ofertę o kursy prowadzone lub przygotowane przez centrum edukacyjne.

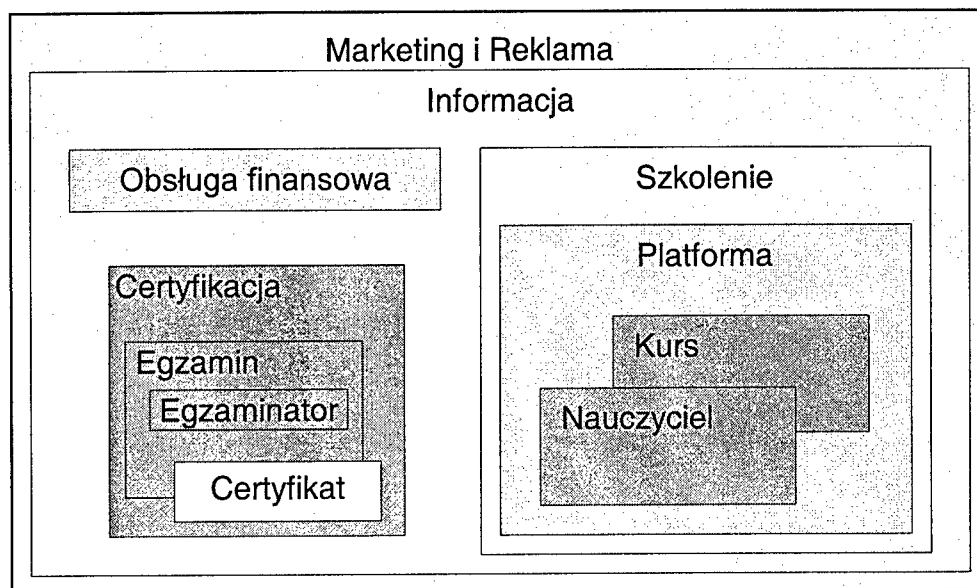
Firma komercyjna: Platforma, Punkt Informacyjny

Centrum edukacyjne: Nauczyciel, Egzamin, Punkt Informacyjny, Kurs

Marketing/Reklama – w zależności od ustaleń prowadzone może być przez firmę, centrum, obie instytucje lub podmiot zewnętrzny.

Certyfikat wystawiany przez firmę, centrum, oba podmioty lub certyfikat zewnętrzny jeżeli firma/centrum/kurs posiada odpowiednie uprawnienia. Dopuszczenie do egzaminu następuje po ukończeniu odpowiedniego kursu oraz uiszczeniu opłat za kurs i egzamin, udokumentowanych potwierdzeniem wpłaty.

Obsługa finansowa prowadzona jest przez komercyjnego partnera, koszt certyfikacji jest wliczony w cenę kursu lub rozliczany całkowicie osobno. Opłaty wpływają na konto komercyjnego partnera, który następnie po zakończeniu Certyfikacji przelewa uzgodnioną kwotę na konto centrum edukacyjnego.



3. TECHNICZNE ASPEKTY MODELU HOSTOWANEGO

Alternatywnym do tradycyjnie budowanej i utrzymywanej infrastruktury technicznej sposobem korzystania ze zintegrowanego systemu informatycznego jest outsourcing / ASP (*Application Service Provider*). Outsourcing zasobów informacyjnych i telekomunikacyjnych (IT) polega na wynajęciu zdalnego dostępu do aplikacji za pośrednictwem łącz telekomunikacyjnych w zamian za okresowo ponoszone opłaty. Na usługę składają się aplikacje, wdrożenie i integracja systemu, dostęp do centrum danych, monitorowanie i ciągła opieka serwisowa. Firma oferująca outsourcing inwestuje w infrastrukturę, konfiguruje ją tak, aby była niezawodna i wydajna i udostępnia gotowe do użytkowania rozwiązanie, zatrudnia wysokokwalifikowany personel, który na bieżąco utrzymuje i aktualizuje system. Dzięki wykorzystaniu najlepszych metod archiwizacji i tworzenia kopii zapasowych oraz zabezpieczeń sieciowych zapewnia najwyższy poziom bezpieczeństwa. Oznacza to, że użytkownik może skoncentrować się na podniesieniu efektywności i wydajności swoich procesów biznesowych i budowaniu przewagi konkurencyjnej. Korzysta przy tym z najnowocześniejszego sprzętu i wysokiej klasy oprogramowania. Nie ponosi kosztów utrzymania własnych służb IT. Nie musi obawiać się kosztów związanych z przeciwdziałaniem starzeniu się infrastruktury informatycznej. Ponadto nie jest ograniczony platformą sprzętową i systemową, jeżeli znajdzie potrzeba zwiększenia liczby licencji na oprogramowanie. Użytkowany przez niego system jest elastyczny i na bieżąco aktualizowany odpowiednio do zmian zachodzących w przepisach. Aktualizacja systemu odbywa się bez zakłócania normalnego trybu pracy firmy. Istotne jest również, że model zakupu w formie outsourcingu daje możliwość zaplanowania w budżecie rocznym niezmienną kwoty na informatyzację.

Platforma sprzętowo-systemowa współpracującej w ramach eksperymentu z Politechniką Gdańską firmy Incenti składa się z następujących zasobów: serwery (SUN lub Intel), systemu operacyjnego Sun Solaris oraz systemu zarządzania procesem kształcenia

iLearning Oracle, bazy danych 9i Oracle oraz serwer aplikacyjny Oracle 9iAS. Praktycznie outsourcing polega na zdalnym dostępie do aplikacji pakietu zintegrowanego, które znajdują się w Centrum Przetwarzania Danych (DC) Incenti, poprzez łącza telekomunikacyjne. DC we Wrocławiu posiada styk z dwoma niezależnymi ISP: TP S.A. – 1Mbps (trakt ATM OC-3) oraz Crowley Data Poland – 2.5Mbps – łącze bezprzewodowe (obecnie w fazie rozbudowy do traktu E-3). Gwarantowana przepustowość łącz jest równa maksymalnej ich pojemności. Dostęp do Internetu realizowany jest z wykorzystaniem protokołu BGP co gwarantuje niezawodność usługi.

4. PODSUMOWANIE - PRZYKŁAD WDROŻENIA SZKOLEŃ FIRMY KOMERCYJNEJ PRZEZ INSTYTUCJĘ EDUKACYJNĄ

Aktualnie przygotowywane jest pilotażowe wdrożenie kursów z zakresu obsługi programu MS EXCEL dla studentów Wydziału Inżynierii Lądowej Politechniki Gdańskiej. Zaproponowano pierwszy wariant modelu „Centrum Szkoleniowe”. Idea tego wdrożenia opiera się na udostępnieniu studentom kursu obsługi MS EXCEL oraz umożliwieniu zainteresowanym studentom przystąpienia do egzaminu, który kwalifikuje do uzyskania certyfikatu potwierdzającego zdobyte wiadomości.

Zaplanowano następujący udział instytucji partnerskich:

Firmy komercyjne (Oracle, Incenti, WiedzaNet): Platforma, Kurs.

Centrum edukacyjne (WIL + CEN PG): Nauczyciel, Egzamin, lokalny Punkt Informacyjny

Do egzaminu student będzie mógł przystąpić po ukończeniu kursu online stopnia 1,2 lub 3. Certyfikat wystawiony będzie przez CEN PG.

BIBLIOGRAFIA

- [1] Grabowska A., Jankowski R., Szymczak Cz.: *Projekt CURE . wykorzystanie technik kształcenia na odległość*. W: Proceedings. Konferencja Technologie Informacyjne. Gdańsk. Polska, 19-21 maja 2003. Ed: Zeszyty naukowe Wydziału ETI Politechniki Gdańskiej, Gdańsk 2003, zeszyt nr 2, s.450-456.
- [2] Singh H., Reed C.: *A White Paper : Achieving Success with Blended Learning*. Centra Software 2001
- [3] Oracle Learning Management :
http://www.oracle.com/applications/human_resources/index.html?learning.html
- [4] Incenti (ASP i E-Learning): <http://www.incenti.pl>
- [5] WiedzaNet : <http://www.wiedzanet.pl>

ORACLE I-LEARNING PLATFORM IN „BLENDED LEARNING” MODELS

Summary

In the article different organisational models for „blended learning” implemented with i-learning Oracle platform are presented. Co-operation between private and public institutions involved in e-learning is discussed. Online Excel training for Civil Engineering students is planned to be implemented by four partners: Oracle [1], [3], Incenti S.A. [4], WiedzaNet Sp. z o.o. [5], CEN PG.

Jerzy Nawrocki, Renata Peliwo, Jacek Maza

Instytut Informatyki, Politechnika Poznańska

INTERNETOWY SYSTEM INFORMACYJNY O PROJEKTACH STUDENCKICH

Streszczenie

W artykule przedstawiony został system realizowany w Instytucie Informatyki Politechniki Poznańskiej w ramach Studia Rozwoju Oprogramowania. Ma on pełnić funkcję informacyjną i wspomagać prace organizacyjne związane z projektami studenckimi. Realizacja systemu jest kontynuacją projektu opisanego w pracy [7]. Powstał wówczas internetowy „*Rejestr Spraw*”. Obecnie nacisk położony jest na komunikację i organizację projektów w początkowej fazie ich realizacji.

1. WSTĘP

Aktywne formy kształcenia wciąż są w cenie. Jedną z takich form jest kształcenie informatyków poprzez organizowanie zajęć na zasadach przypominających funkcjonowanie małej firmy informatycznej. Na Politechnice Poznańskiej tego typu zajęcia są prowadzone od 1998 roku pod nazwą Studia Rozwoju Oprogramowania. Każdego roku jest realizowanych 11-12 rocznych przedsięwzięć informatycznych dla klientów zewnętrznych (są wśród nich nawet klienci ze Szwecji). Dobre zarządzanie kilkunastoma projektami nie jest zadaniem trywialnym. Jest to tym trudniejsze, że warunki akademickie wymuszają coroczną wymianę pracowników (kończy się rok akademicki i studenci odchodzą z zajmowanych przez nich stanowisk). Kluczowym czynnikiem warunkującym powodzenie jest dobra komunikacja i właściwy przepływ informacji. Aby przedsięwzięcia zakończyły się sukcesem potrzebne są dobre rozwiązania organizacyjne i wsparcie narzędziowe ukierunkowane na zagadnienia komunikacji.

Celem artykułu jest przedstawienie internetowego systemu informacyjnego przeznaczonego do gromadzenia i prezentacji danych o projektach studenckich. System jest budowany w Instytucie Informatyki Politechniki Poznańskiej i jest ukierunkowany na potrzeby Studia Rozwoju Oprogramowania. Wydaje się, że system ten mógłby być stosunkowo łatwo przystosowany do potrzeb innych uczelni zamierzających wprowadzić aktywne formy kształcenia na zasadach podobnych do Studia Rozwoju Oprogramowania.

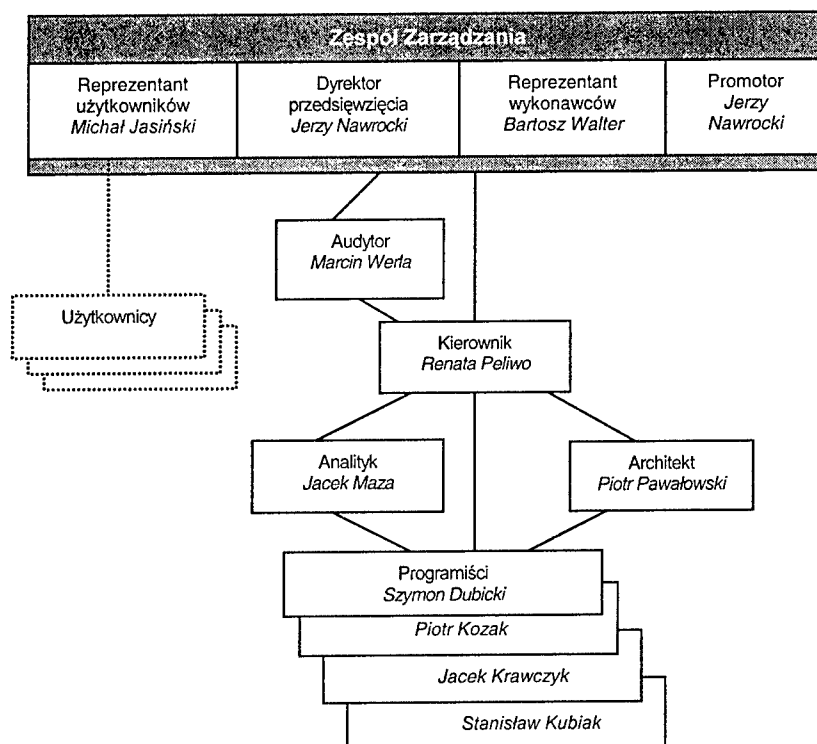
W następnym rozdziale omówiono organizację Studia Rozwoju Oprogramowania. W rozdziale 3. przedstawiono motywacje i uwarunkowania biznesowe dotyczące budowanego systemu. Specyfikację wymagań postanowiono oprzeć na stosunkowo nowej technice zwanej przypadkami użycia [5]. Rozdział 4. poświęcono opisowi systemu oraz

przedstawiono przykładowe przypadki użycia. W rozdziale 5. przedstawiono zarys architektury systemu, a w rozdziale 6. opisano krótko zarządzanie realizacją tego przedsięwzięcia.

2. STUDIO ROZWOJU OPROGRAMOWANIA

2.1 Organizacja Studia Rozwoju Oprogramowania

Studio Rozwoju Oprogramowania działa na zasadach przypominających funkcjonowanie małej firmy informatycznej. Jest ono prowadzone na Politechnice Poznańskiej od 1998 roku. W ramach Studia Rozwoju Oprogramowania realizowanych jest 11-12 rocznych przedsięwzięć informatycznych dla klientów (firmy i instytucje). Łącznie bierze w nich udział ok. 120 osób. Studenci podzieleni są na dwie grupy, z których każda ma opiekuna pełniącego rolę kierownictwa Studia.



Rys.1. Przykładowa organizacja zespołu (projekt e-SDS).

W realizacji każdego przedsięwzięcia bierze udział 12 osób, w tym opiekun studium, klient, promotor, reprezentant wykonawców, reprezentant użytkowników oraz studenci III, IV i V roku. Skład zespołu realizującego przedsięwzięcie jest zbliżony do składu zespołu metodyki PRINCE 2 [1]. Wzbogacono go o funkcję promotora, analityka i architekta oprogramowania. Funkcja promotora jest niezbędna ze względu na akademicki charakter

przedsięwzięcia. Dla studentów III roku realizowany projekt jest pracą inżynierską. Ich zadaniem jest projektowanie, implementacja, testowanie oraz dokumentowanie systemu. Studenci IV roku pełnią rolę kierownika oraz analityka. Studenci V roku są odpowiedzialni za kontrolę jakości oraz architekturę systemu.

Skład i organizacja zespołu została przedstawiona na rys. 1 na przykładzie zespołu realizującego to przedsięwzięcie.

Praca Studia Rozwoju Oprogramowania odbywa się w cyklu rocznym. W czerwcu (pod koniec III roku studiów) studenci dokonują wyboru specjalizacji. Do końca września klienci zgłaszają propozycje przedsięwzięć. Na początku października studenci IV roku wybierają spośród nadesłanych propozycji przedsięwzięcia, które chcą realizować. Znajdują oni promotora dla studentów III roku oraz dobierają się w zespoły ze studentami V roku. W listopadzie do zespołu dołączają studenci III roku. Przedsięwzięcia kończą się najczęściej w czerwcu. Wówczas następuje również obrona pracy inżynierskiej studentów III roku.

2.2. Rola Studia Rozwoju Oprogramowania

Udział w przedsięwzięciach realizowanych w ramach Studia Rozwoju Oprogramowania daje studentom możliwość zdobycia doświadczenia w zakresie prowadzenia przedsięwzięć informatycznych. Sposób realizacji projektów jest odmienny niż projektów realizowanych w ramach innych zajęć na uczelni. Studenci uczą się pracy w większej grupie oraz tego jak właściwie zarządzać przedsięwzięciem. Bardzo cenna jest możliwość współpracy z klientami (firmy oraz instytucje). W przeciwieństwie do większości projektów realizowanych na zajęciach, stworzone w ramach Studia oprogramowanie jest wdrażane oraz wykorzystywane w różnych organizacjach.

Studio Rozwoju Oprogramowania działa na zasadach małej firmy informatycznej, w związku z czym kluczową rolę odgrywają klienci. Przedstawiciel klienta zgłasza propozycję przedsięwzięcia i otrzymuje potwierdzenie jego realizacji bądź informację o odrzuceniu propozycji (w przypadku gdy nie została wybrana przez studentów). W pierwszym przypadku wchodzi w skład zespołu realizującego przedsięwzięcie, bierze udział w formułowaniu wymagań oraz zatwierdza kolejne wersje systemu. Z punktu widzenia klienta, rola Studia Rozwoju Oprogramowania polega na dostarczeniu oraz wdrożeniu zamówionego oprogramowania.

3. MOTYWACJA I UWARUNKOWANIA BIZNESOWE

3.1. Motywacja

Dobre zarządzanie organizacją programistyczną wymaga dobrej komunikacji. Niestety w przeszłości zdarzało się, że reprezentant klienta nie był na czas poinformowany, że studenci nie wybrali jego propozycji przedsięwzięcia lub też promotor dopytywał się o status przedsięwzięcia (najczęściej miało to miejsce na początku roku akademickiego). Ilość danych związanych z każdym projektem jest bardzo duża. Pojawia się problem ich przechowywania, przetwarzania oraz przekazywania wszystkim zainteresowanym osobom. Zdarzało się, że studenci nie znali wszystkich propozycji, spośród których mogą wybrać przedsięwzięcie.

Rozwiązaniem tego typu problemów jest internetowy system informacyjny o projektach studenckich. Głównym celem systemu jest poprawa komunikacji oraz organizacji

projektów. System ma wspomagać zgłaszanie propozycji projektów oraz sprawy organizacyjne związane z zapisami. Każda osoba zainteresowana projektem będzie mogła obserwować stopień jego zaawansowania poprzez Internet. Dane przechowywane w systemie będą łatwo dostępne. Każda osoba może je uzupełnić lub zmienić, dzięki czemu będą zawsze aktualne. System będzie automatycznie informował zainteresowane osoby o zmianach dotyczących projektów oraz o nowych propozycjach. Informacje będą przechowywane, grupowane i prezentowane w postaci stron internetowych. Z systemu korzystać będą wszystkie osoby związane z realizacją projektów.

3.2. Uwarunkowania biznesowe

Pierwszy kontakt potencjalnego klienta ze Studium Rozwoju Oprogramowania jest bardzo ważny. Poczucie braku organizacji może spowodować, że zrezygnuje on z chęci realizacji przedsięwzięcia w ramach Studia lub też ewentualne złe wrażenia na początku będą rzutować na ogólne wrażenie aż do końca przedsięwzięcia (nawet jeśli będzie ono bardzo dobrze zarządzane w dalszych etapach). Poczucie zagubienia może również niekorzystnie wpłynąć na stosunek studentów do realizacji przedsięwzięcia. Warto zainwestować w zbudowanie systemu, który uporządkuje wstępną fazę realizacji przedsięwzięć. Odbiór Studia Rozwoju Oprogramowania będzie dużo lepszy.

Ważnym aspektem jest pracochłonność koordynacji działań we wstępnej fazie realizacji przedsięwzięć. Wymagane jest bardzo duże zaangażowanie ze strony kierownictwa Studia. Najwięcej czasu zajmują drobne kwestie organizacyjne, które mogą zostać zautomatyzowane. Realizacja systemu pozwoli skupić uwagę kierownictwa na bardziej złożonych sprawach. Po wdrożeniu systemu czas poświęcany przez opiekunów Studia powinien się zmniejszyć.

Kolejnym zyskiem z wdrożenia systemu jest przyspieszenie rozpoczęcia prac nad realizacją przedsięwzięć. Mogą one szybciej wejść w etap inżynierii wymagań. Obecnie formowanie grup i wybór przedsięwzięć odbywa się w połowie października. Oczekuje się, że dzięki korzystaniu z systemu większość przedsięwzięć będzie kompletnych już na początku października, co powinno dać przyspieszenie w realizacji każdego z przedsięwzięć przynajmniej o tydzień.

4. CECHY FUNKCJONALNE SYSTEMU

Wymagania funkcjonalne dla tworzonego systemu zostały opisane w postaci przypadków użycia [5]. Stanowią one podstawę do projektowania oraz implementacji przedstawianego systemu.

Przypadek użycia opisuje interakcję aktora (użytkownika lub sprzętu) z systemem za pomocą kroków, prowadzących do osiągnięcia określonego celu. W zależności od potrzeb każdy przypadek może być opisany na różnym poziomie szczegółowości. Poziomy: biznesowy, użytkownika oraz funkcji systemowych dostarczają coraz bardziej szczegółowych informacji. Z każdym przypadkiem mogą być związane pewne atrybuty, jak np. wartość dla klienta, koszt realizacji, warunki wstępne, cel. Liczba atrybutów jest różna w zależności od poziomu szczegółowości na jakim jest opisany przypadek.

Specyfikacja wymagań funkcjonalnych dla realizowanego systemu zajmuje ok. 20 stron. W związku z tym opis wymagań w postaci przypadków użycia zostanie przedstawiony na przykładzie. Pozostałe wymagania zostaną opisane mniej szczegółowo.

4.1. Podstawowa funkcjonalność systemu

System ułatwia zbieranie i prezentację informacji o przedsięwzięciach realizowanych w ramach Studia Rozwoju Oprogramowania. W zależności od pełnionej roli (patrz rozdział 2) użytkownik może wykonać różne czynności. Potencjalni klienci rejestrują się w systemie. Mają możliwość zapoznania się z organizacją Studia Rozwoju Oprogramowania. Za pomocą systemu mogą oni zgłaszać propozycje przedsięwzięć, które chcieliby by były zrealizowane. Promotorzy i studenci mogą zapoznać się ze zgłoszonymi propozycjami oraz wybrać tę, którą chcą realizować. W przypadku kadry Studia Rozwoju Oprogramowania system wspomaga kontrolę realizacji przedsięwzięć oraz ułatwia zarządzanie nimi. System powinien automatycznie informować zainteresowane osoby o zmianach w przedsięwzięciu oraz o nowych propozycjach. W tym celu wysyłane są wiadomości (e-mail) do zainteresowanych osób.

4.2. Przypadki użycia

Przykładowy opis dotyczy roli Studia Rozwoju Oprogramowania z punktu widzenia klienta. Aby uzyskać oprogramowanie klient musi zgłosić propozycję projektu. Wcześniej musi zostać zarejestrowana organizacja oraz jej przedstawiciel.

Tablica 4.1

Przykładowy przypadek użycia – poziom biznesowy

Nazwa	Uzyskaj oprogramowanie
Aktor	Przedstawiciel klienta
Cel	Uzyskanie oprogramowania wytworzonego w ramach Studia Rozwoju Oprogramowania

Tablica 4.2

Przykładowe przypadki użycia – poziom użytkownika

Nazwa	Zgłoś propozycję projektu
Aktor	Przedstawiciel klienta
Cel	Rejestracja propozycji projektu w systemie
Warunek wstępny	Przedstawiciel klienta musi być zalogowany
Zdarzenie inicjujące	Przedstawiciel klienta chce zgłosić propozycję projektu
Scenariusz główny	1. System wyświetla formularz do wprowadzenia danych. 2. Przedstawiciel klienta wprowadza dane i je potwierdza. 3. System aktualizuje dane i potwierdza zgłoszenie. 4. Przedstawiciel klienta kończy proces zgłaszania propozycji projektu.
Alternatywne zachowania	Brak
Nazwa	Zarejestruj organizację

Aktor	Przedstawiciel klienta
Cel	Zarejestrowanie organizacji w systemie
Warunek wstępny	Brak
Zdarzenie inicjujące	Przedstawiciel klienta chce zarejestrować organizację, jako potencjalnego klienta Studia Rozwoju Oprogramowania
Scenariusz główny	1. System wyświetla formularz do wprowadzenia danych. 2. Przedstawiciel klienta wprowadza dane i je potwierdza. 3. System aktualizuje dane i potwierdza rejestrację. 4. Przedstawiciel klienta kończy proces rejestracji
Alternatywne zachowania	3a. Organizacja jest już zarejestrowana w systemie. 3a.1. System informuje, że dana organizacja jest już zarejestrowana w systemie. Przejsście do kroku 4.

5. ZARYS ARCHITEKTURY

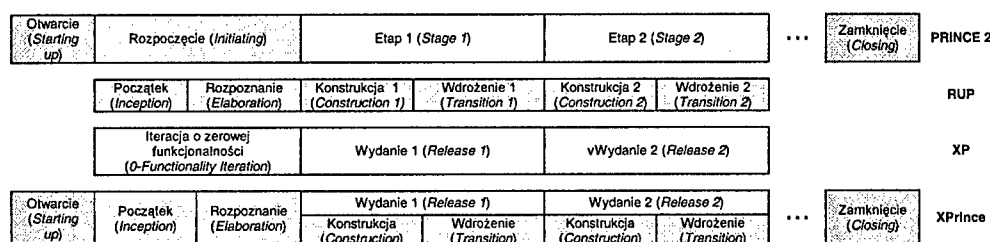
Wydaje się, że właściwym rozwiązaniem będzie zastosowanie architektury trójwarstwowej: klient – serwer – baza danych. Aby uzyskać dostęp do zasobów systemu klient (przeglądarka WWW) łączy się z serwerem WWW. Serwer ten przetwarza żądanie w oparciu o dane uzyskane z serwera realizowanego systemu (e-SDS) i odsyła klientowi odpowiedź w postaci strony WWW. Serwer e-SDS łączy się z bazą danych. Korzysta również z serwera poczty elektronicznej w celu wysyłania wiadomości (e-mail) do użytkowników systemu. Celem realizowanego przedsięwzięcia jest stworzenie komponentów serwera e-SDS oraz struktury bazy danych.

Wszystkie komponenty serwera są tworzone przy pomocy TDK (Turbine Development Kit) [3]. Jest to szkielet aplikacji internetowej bazujący na Servletach Javy. TDK składa się z trzech części: Turbine, Velocity oraz Torque. Velocity jest wykorzystywane dla oddzielenia warstwy logiki od warstwy prezentacji. Wzorcowe pliki HTML wypełniane są danymi z warstwy logiki. Torque stanowi warstwę pośredniczącą pomiędzy relacyjną bazą danych a aplikacją w Javie. Umożliwia definicję schematu bazy danych w pliku XML, stworzenie tej bazy oraz wygenerowanie potrzebnych do jej obsługi klas. Zapewnia to dostęp do danych z poziomu aplikacji. Dzięki temu dane z relacyjnej bazy danych są od razu dostępne jako obiekty aplikacji i programista nie musi manualnie zarządzać bazą danych. Turbine przystosowane jest do współpracy z wieloma bazami danych. W realizowanym systemie użyta zostanie baza danych MySQL [4]. Została ona przetestowana przez twórców Turbine i gwarantuje poprawne działanie oraz wystarczającą efektywność.

6. ZARZĄDZANIE REALIZACJĄ PRZEDSIĘWZIĘCIA

Właściwe zarządzanie przedsięwzięciem informatycznym jest głównym czynnikiem sukcesu. Przedsięwzięcie prowadzone jest zgodnie z metodyką XPrince bazującą na PRINCE 2 (*Projects In Controlled Environments*) [1], RUP (*Rational Unified Process*) [2] oraz XP (*Extreme Programming*) [6]. Relacje między XPrince a wspomnianymi wcześniej metodykami przedstawiono na rysunku 2. Przedsięwzięcie zostało podzielone na etapy. Są

dwa wydania systemu, z których każde zostało podzielone na dwa przyrosty. Poszczególne etapy przedsięwzięcia oraz wykonywane w nich czynności stanowią połączenie etapów metodyk PRINCE 2 oraz RUP. Programiści pracują zgodnie z zasadami metodyki programowania XP. Więcej informacji na temat metodyki XPrince można znaleźć w pracy [8].



Rys.2. XPrince.

7. ZAKOŃCZENIE

W artykule przedstawiono internetowy system informacyjny przeznaczony do gromadzenia i prezentacji informacji dotyczących projektów studenckich. System budowany jest w Instytucie Informatyki Politechniki Poznańskiej. Ma on wspomagać prace organizacyjne związane z funkcjonowaniem Studia Rozwoju Oprogramowania. Na podstawie kilkuletniego doświadczenia (projekty realizowane są od 1998 roku) widoczna jest potrzeba realizacji opisanego systemu. Wydaje się, że system poprawi odbiór Studia Rozwoju Oprogramowania wśród studentów oraz potencjalnych klientów. Obecnie zgłaszanie propozycji i początkowy etap realizacji przedsięwzięć wymagają częstych kontaktów klienta z kierownikami Studia. Klient może nie być zainteresowany tak częstą komunikacją. Realizacja systemu pozwoli mu obserwować status przedsięwzięcia w dowolnym czasie i miejscu. Korzystanie z systemu pozwoli zaoszczędzić czas przeznaczany obecnie na sprawy organizacyjne, które mogą zostać zautomatyzowane. Dzięki temu więcej czasu pozostaje na planowanie i realizację przedsięwzięcia.

PODZIĘKOWANIA

Autorzy dziękują Szymonowi Dubickiemu, Piotrowi Kozakowi, Jackowi Krawczykowi, Stanisławowi Kubiakowi za prace programistyczne, Piotrowi Pawałowskiemu i Marcinowi Werli za pełnienie roli architekta i kontrolera jakości oraz Bartkowi Walterowi za nadzór nad realizacją przedsięwzięcia.

Praca została wykonana w ramach grantu BW-91-399/04.

BIBLIOGRAFIA

- [1] *Managing Successful Projects with PRINCE 2*, CCTA, The Stationary Office, Norwich, 2002.
- [2] P. Kroll, P. Kruchten: *The Rational Unified Process Made Easy*, Addison-Wesley, 2003.
- [3] TDK <http://jakarta.apache.org/turbine>
- [4] MySQL <http://www.mysql.com>
- [5] Steve Adolph, Paul Bramble: *Patterns for Effective Use Cases*, Addison-Wesley, 2003.
- [6] Kent Back: *Extreme Programming: Embrace Change*, Addison-Wesley, 2000.

- [7] J. Nawrocki, P. Pawałowski i K. Pospiech: *System Internetowy wspomagający wytwarzanie oprogramowania metodą PRINCE2*, Zeszyty naukowe Politechniki Gdańskiej nr 1, 2003.
- [8] Jerzy Nawrocki, Michał Jasiński: *XPrince: Extreme Programming in Controlled Environment*, Politechnika Poznańska, Poznań 2004.

INTERNET-BASED SYSTEM SUPPORTING STUDENTS PROJECTS

Summary

The article contains the description of the system that is developed at Institute of Computer Science at Poznań University of Technology. It develops for needs of Software Development Studio (SDS). SDS is a module of the Master in Software Engineering Programme. It can be compared to a small computer company. Each year 11-12 one-year projects are performed for outside customers. At each project take part about 12 people.

The main factor that determines the success is good communication and good flow of information. Good organisational solutions and tools support are required to achieve a success. The system is going to perform information function, the main aim of it is to improve communication and organisation of the projects.

Introduction of the system will let save time that is now used to solve organizational matters. Therefore, more time will be saved to plan and perform the project.